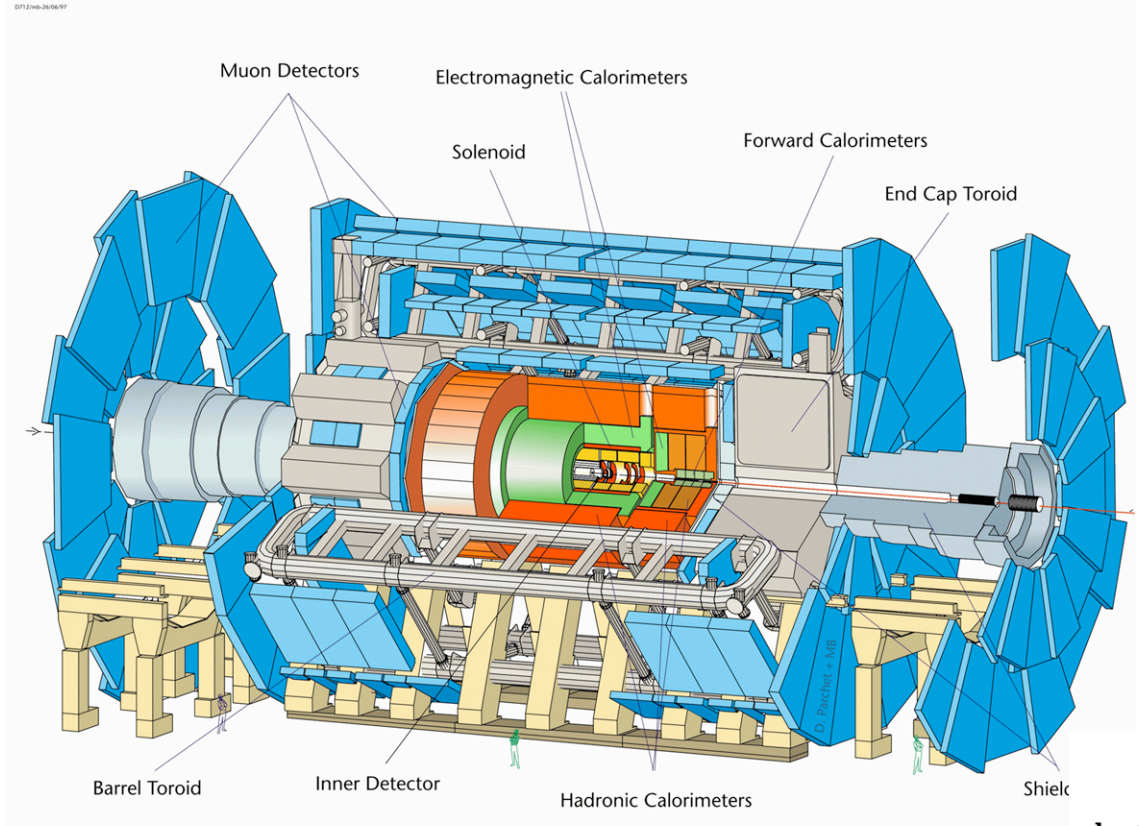
ATLAS Software and Computing - Are we ready for first physics?

David Quarrie
LBNL
drquarrie@lbl.gov

Introduction

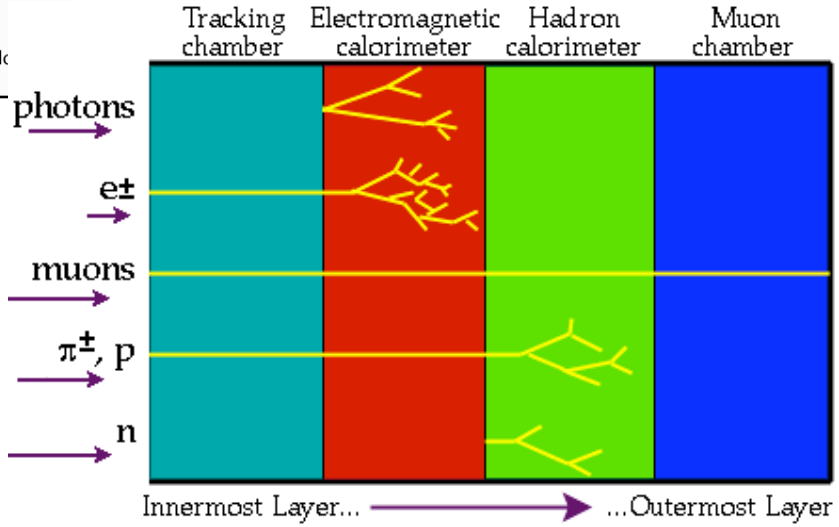
- One word answer - No!
  - Real data will never look quite like one predicts and we have to anticipate the unexpected
- Longer answer:
  - What we are doing to try to minimize the unexpected and then deal with it
  - Brief discussion of ATLAS Computing Model
  - Strategy for software releases and validation
  - Strategy to cope with problems when first real physics data is encountered
  - Computing and Software Stress tests
- Schedule
  - Beampipe closed in August
  - First collisions in October

ATLAS



Length : ~ 46 m
Radius : ~ 12 m
Weight : ~ 7000 tons
~ 10^8 electronic channels
~ 3000 km of cables

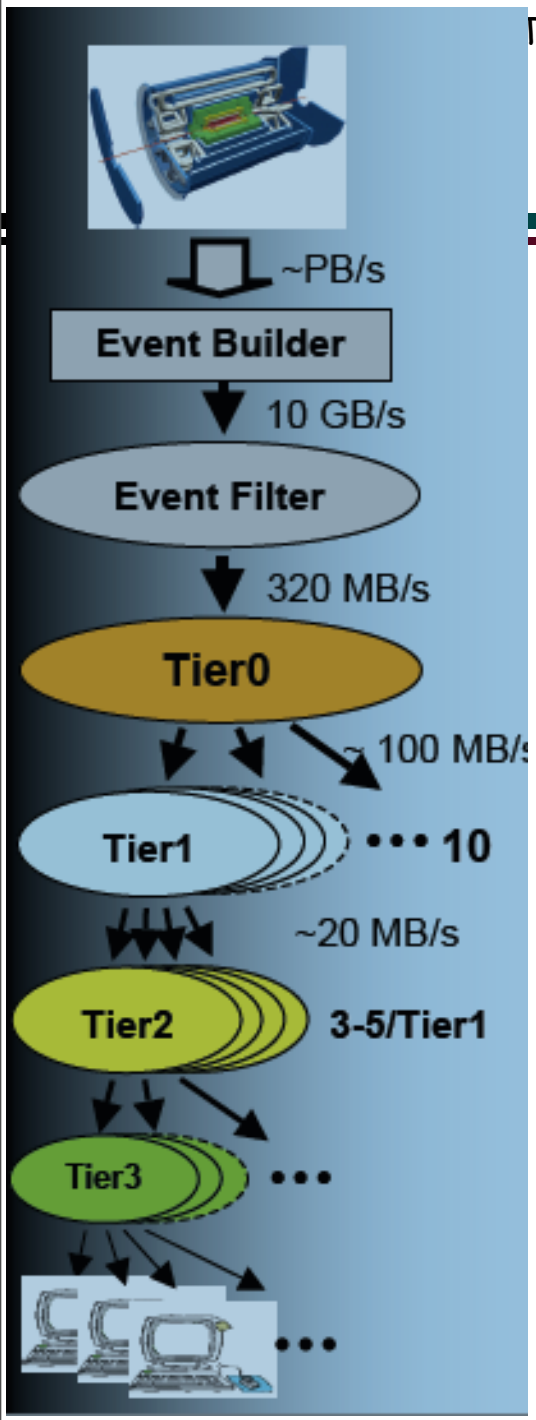
- **Tracking ($|\eta|<2.5$, B=2T) :**
  - **-- Si pixels and strips**
  - **-- Transition Radiation Detector (e/π separation)**
- **Calorimetry ($|\eta|<5$) :**
  - **-- EM : Pb-LAr**
  - **-- HAD: Fe/scintillator (central), Cu/W-LAr (fwd)**
- **Muon Spectrometer ($|\eta|<2.7$) :**
  - **air-core toroids with muon chambers**



ATLAS Computing Characteristics

- Large, complex detector
  - ~10^8 channels
- Long lifetime
  - Project started in 1992, first data in 2008, last data >2030?
- 320 MB/sec raw data rate
  - ~3 PB/year
- Large, geographically dispersed collaboration
  - ~2k scientific authors, 167 institutions, 36 countries
  - Many are, or will become, software developers even just for physics analysis
- Scale and complexity reflected in software
  - >1000 packages, >8000 C++ classes, >5M lines of code
  - ~70% code is algorithmic (written by physicists)
  - ~30% infrastructure, framework (written by primarily by sw engineers)

ATLAS Computing Model



- TDAQ
  - At Point1 on LHC ring
- Tier0
  - CERN Computing Center
- 10 Tier1s
  - BNL (Brookhaven, US), NIKHEF/SARA (Amsterdam, NL), CC-IN2P3 (Lyon, FR), FZK (Karlsruhe, DE), NDGF (DK/SE/NO), PIC (Barcelona, ES), CNAF (Bologna, IT), RAL (Chilton, UK), TRIUMF (Vancouver, CA), ASGC (Taipei, TW)
- ~37 Tier2s, in most participating countries
- ~170 Tier3s in all participating institutions
- Data distribution and access is critical

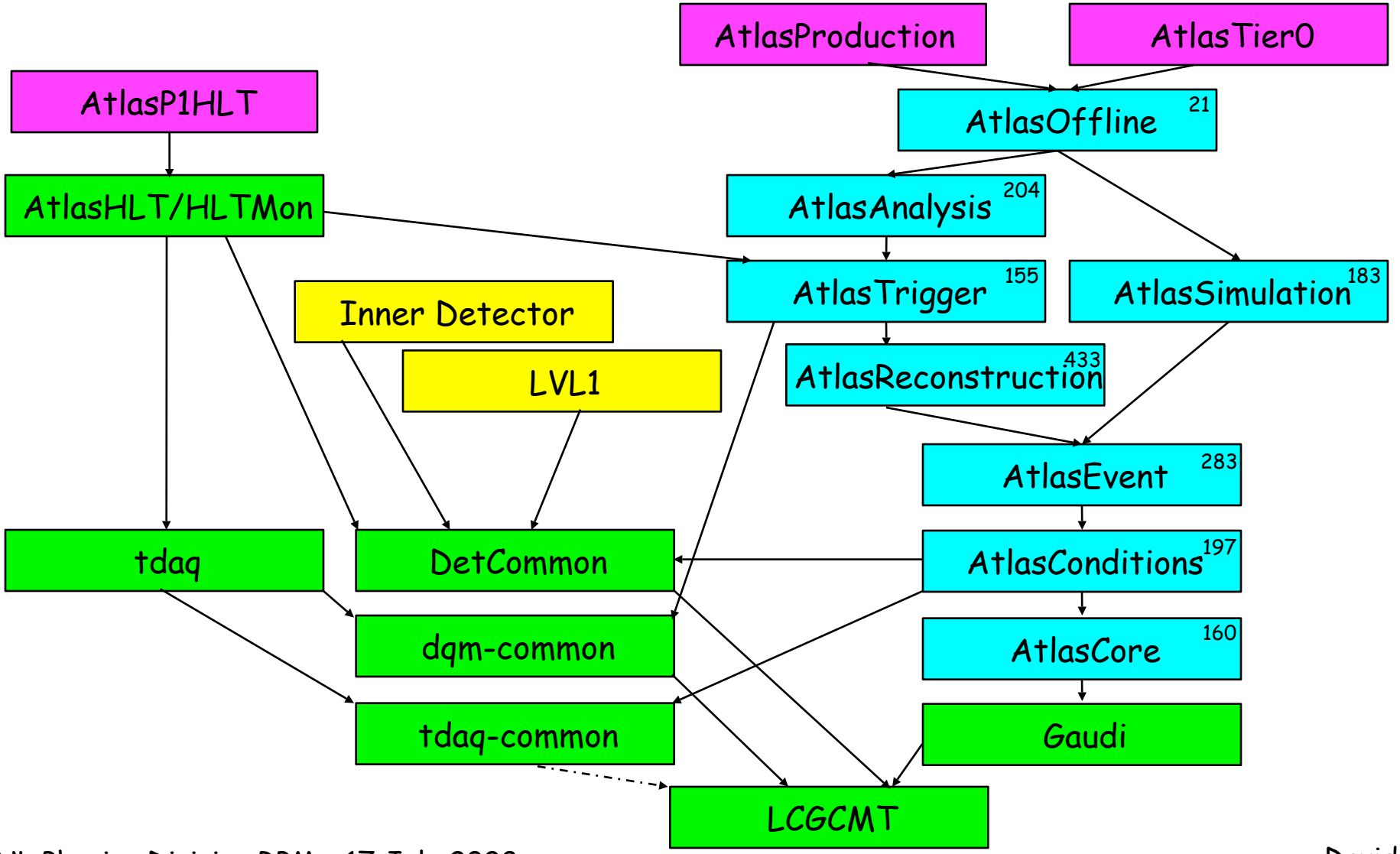
Data Replication and Distribution

- RAW: ~1.6MB/event
  - Original data at Tier0
  - Complete replica distributed among all Tier1s
    - Randomized datasets to make reprocessing more efficient
- Event Summary Data (ESD): ~1MB/event
  - Produced by primary reconstruction at Tier0; exported to 2 Tier1s
  - Subsequent versions from re-reprocessing at Tier1s from their RAW data are stored locally and replicated to another Tier1 (2 copies total)
- Analysis Object Data (AOD): ~200kB/event
  - Completely replicated at each Tier1
  - Partially replicated to Tier2s (30%-50% in each Tier2) so complete set for each Tier1
- Derived Physics Data (DPD): ~10kb/event
  - Produced at Tier1s and Tier2s for physics groups
- TAG: ~1kB/event
  - TAG databases replicated to all Tier1s (as Oracle DB and ROOT files)

Software Structure

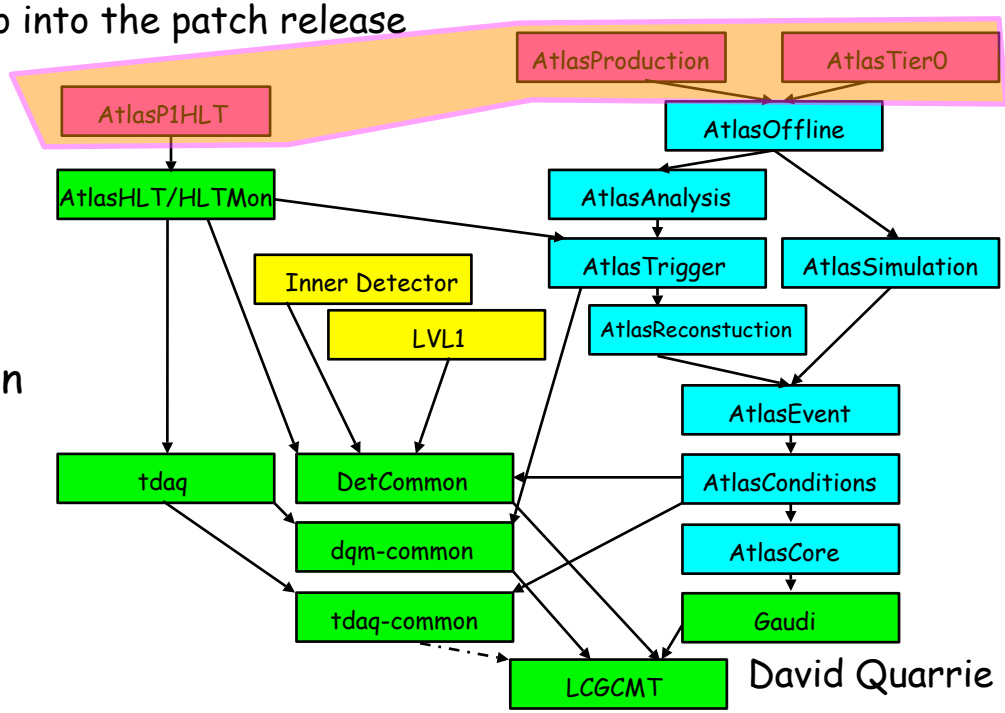
- Primarily C++ or Python classes
  - Some FORTRAN code (mainly generators but some muon and magnetic field)
- Packages
  - Management Unit
  - Container or Leaf
    - Container just corresponds to a directory (no other content allowed) but management unit
    - Leaf contains source code and/or scripts
    - Some leaf packages act as "glue" to external software packages
  - Vary as function of time - version as a "snapshot"
- Projects
  - Groups of packages having similar dependencies built as a unit
  - Version is snapshot of fixed package versions
- Releases
  - Snapshot of complete project hierarchy

Project Dependencies



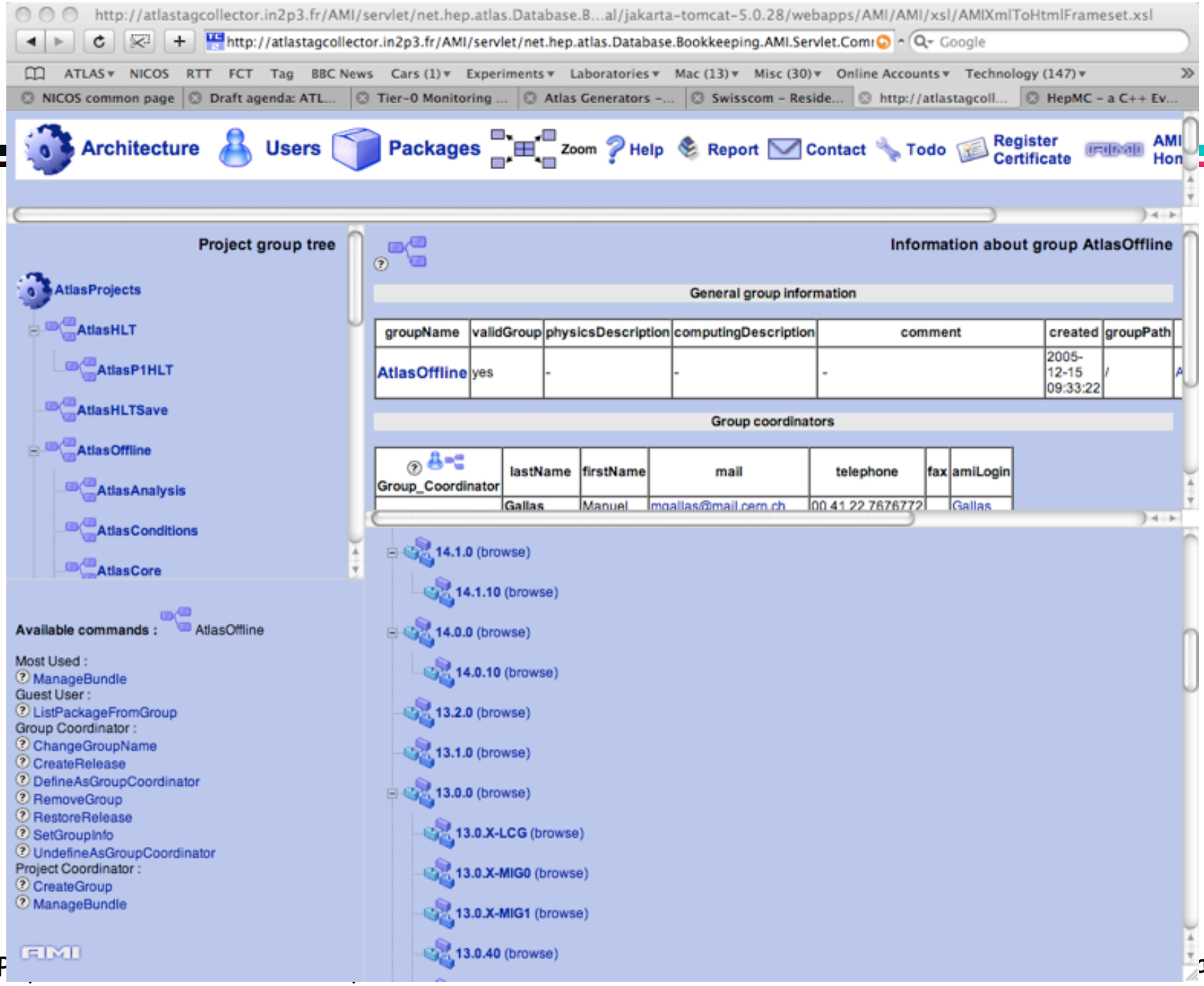
Patch Projects

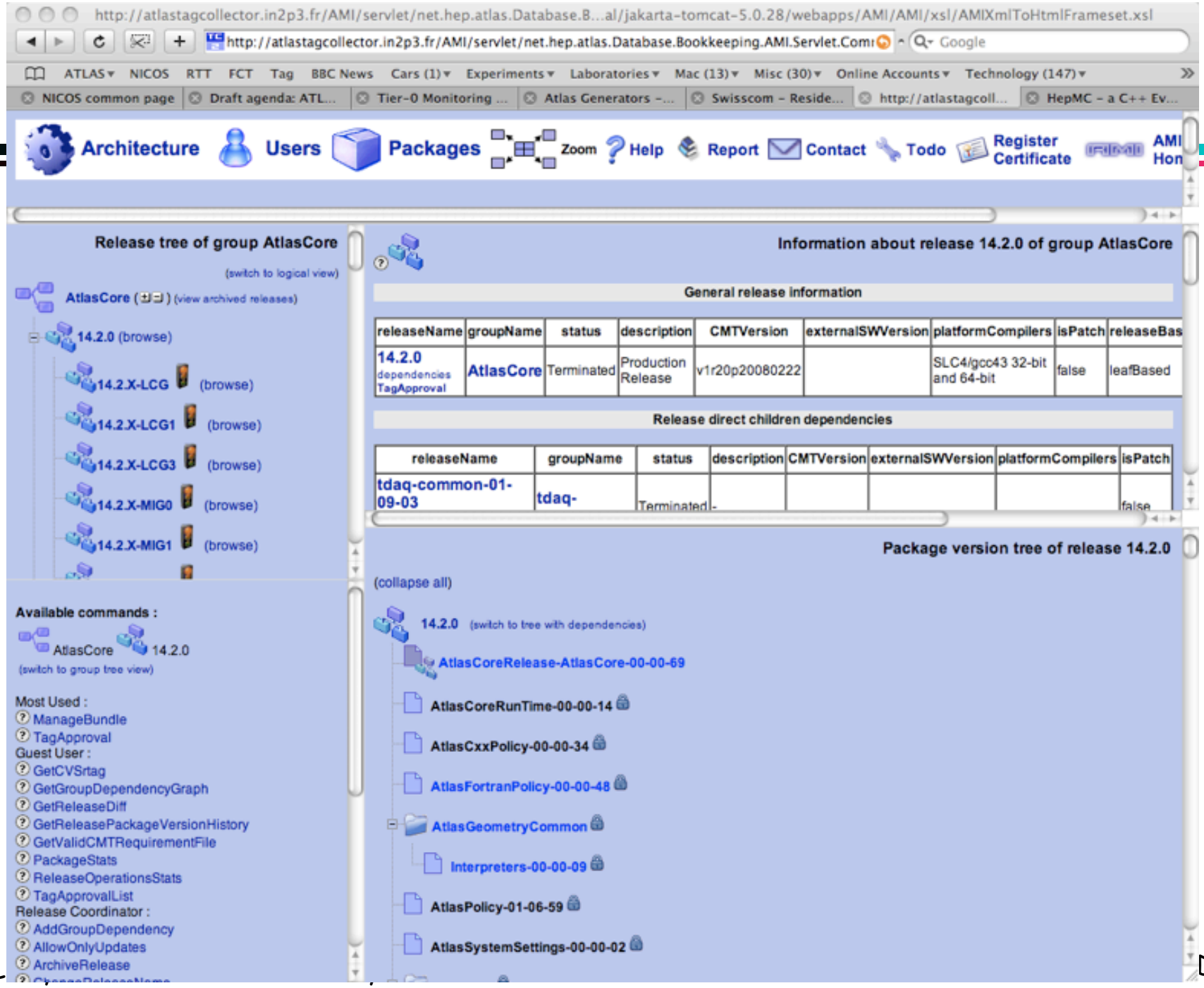
- Special projects that sit at top of project hierarchy
- Prepend to various paths (PATH, LD_LIBRARY_PATH, PYTHONPATH, etc.)
- Contain package override versions
- Patches restricted to Python, scripts, and C++ .cxx files
  - Unless a C++ .h file is private to the package or has only a few client packages
    - In the latter case the clients must also go into the patch release
- Multiple patch projects supported - e.g.
  - One for Tier0 fixes requiring rapid turn-around and reduced validation
  - One for GRID deployment needing less rapid turn-around but increased validation



Release Strategy

- Plan for controlled and validated addition of high priority and late deliverables
  - Establish mechanisms for establishing priorities and dealing with late deliverables
- Plan for rapid (but validated) response to problems with first physics data
  - Extensive cosmics tests; Full Dress Rehearsal (FDR)
  - But first physics data will expose deficiencies
- Staged, incremental releases (~every month)
  - With short branches for bug fixes and patches
- Longer term development moved off to side ("migration" nightlies)
- Baseline release 2-3 months prior to first physics data
  - Emphasis on robustness and technical performance (cpu/memory)
- Strategy replies upon strengthened release coordination and validation
  - Frequent meetings to review high priority deliverables and schedule
  - Validation of one release builds upon previous ones


http://atlastagcollector.in2p3.fr/AMI/servlet/net.hep.atlas.Database.B...al/jakarta-tomcat-5.0.28/webapps/AMI/AMI/xsl/AMIXmlToHtmlFrameset.xsl
http://atlastagcollector.in2p3.fr/AMI/servlet/net.hep.atlas.Database.Bookkeeping.AMI.Servlet.Com
Google
ATLAS NICOS RTT FCT Tag BBC News Cars (1) Experiments Laboratories Mac (13) Misc (30) Online Accounts Technology (147)
NICOS common page
Draft agenda: ATL...
Tier-0 Monitoring ...
Atlas Generators -...
Swisscom - Reside...
http://atlastagcoll...
HepMC - a C++ Ev...
Architecture
Users
Packages
Zoom
Help
Report
Contact
Todo
Register Certificate
AMI Hon
Project group tree
AtlasProjects
AtlasHLT
AtlasP1HLT
AtlasHLTSave
AtlasOffline
AtlasAnalysis
AtlasConditions
AtlasCore
Information about group AtlasOffline
General group information
groupName
validGroup
physicsDescription
computingDescription
comment
created
groupPath
AtlasOffline
yes
-
-
-
2005-12-15 09:33:22
/
Group coordinators
Group_Coordinator
lastName
firstName
mail
telephone
fax
amiLogin
Gallas
Manuel
mgallas@mail.cern.ch
00 41 22 7676772
Gallas
14.1.0 (browse)
14.1.10 (browse)
14.0.0 (browse)
14.0.10 (browse)
13.2.0 (browse)
13.1.0 (browse)
13.0.0 (browse)
13.0.X-LCG (browse)
13.0.X-MIG0 (browse)
13.0.X-MIG1 (browse)
13.0.40 (browse)
Available commands : AtlasOffline
Most Used :
ManageBundle
Guest User :
ListPackageFromGroup
Group Coordinator :
ChangeGroupName
CreateRelease
DefineAsGroupCoordinator
RemoveGroup
RestoreRelease
SetGroupInfo
UndefineAsGroupCoordinator
Project Coordinator :
CreateGroup
ManageBundle


http://atlastagcollector.in2p3.fr/AMI/servlet/net.hep.atlas.Database.B...al/jakarta-tomcat-5.0.28/webapps/AMI/AMI/xsl/AMIXmlToHtmlFrameset.xsl
http://atlastagcollector.in2p3.fr/AMI/servlet/net.hep.atlas.Database.Bookkeeping.AMI.Servlet.Com
Google
ATLAS NICOS RTT FCT Tag BBC News Cars (1) Experiments Laboratories Mac (13) Misc (30) Online Accounts Technology (147)
NICOS common page
Draft agenda: ATL...
Tier-0 Monitoring ...
Atlas Generators -...
Swisscom - Reside...
http://atlastagcoll...
HepMC - a C++ Ev...
Architecture
Users
Packages
Zoom
Help
Report
Contact
Todo
Register Certificate
AMI Hon
Release tree of group AtlasCore
(switch to logical view)
AtlasCore (view archived releases)
14.2.0 (browse)
14.2.X-LCG (browse)
14.2.X-LCG1 (browse)
14.2.X-LCG3 (browse)
14.2.X-MIG0 (browse)
14.2.X-MIG1 (browse)
Available commands :
AtlasCore 14.2.0
(switch to group tree view)
Most Used :
ManageBundle
TagApproval
Guest User :
GetCVSrtag
GetGroupDependencyGraph
GetReleaseDiff
GetReleasePackageVersionHistory
GetValidCMTRequirementFile
PackageStats
ReleaseOperationsStats
TagApprovalList
Release Coordinator :
AddGroupDependency
AllowOnlyUpdates
ArchiveRelease
Information about release 14.2.0 of group AtlasCore
General release information
releaseName | groupName | status | description | CMTVersion | externalSWVersion | platformCompilers | isPatch | releaseBas
14.2.0 dependencies TagApproval | AtlasCore | Terminated | Production Release | v1r20p20080222 | | SLC4/gcc43 32-bit and 64-bit | false | leafBased
Release direct children dependencies
releaseName | groupName | status | description | CMTVersion | externalSWVersion | platformCompilers | isPatch
tdaq-common-01-09-03 | tdaq- | Terminated | | | | | false
Package version tree of release 14.2.0
(collapse all)
14.2.0 (switch to tree with dependencies)
AtlasCoreRelease-AtlasCore-00-00-69
AtlasCoreRunTime-00-00-14
AtlasCxxPolicy-00-00-34
AtlasFortranPolicy-00-00-48
AtlasGeometryCommon
Interpreters-00-00-09
AtlasPolicy-01-06-59
AtlasSystemSettings-00-00-02

NICOS - NIghtly COntrol System

- Nightly build system runs on suite of lxbuild machines at CERN
- Several different branches open simultaneously
- Several different platforms
  - Although currently SLC4 is the only production validated platform
- Multiple multi-core machines allow parallelism
  - Different branches run on different machines
  - Different platforms within same branch run on different machines and merged on AFS
  - Some projects have no cross-dependencies and could be built in parallel on multi-core machine (not actually doing this at the moment since gain not worth it)
  - Tbroadcast provides package level parallelism within project
    - Packages within project with no cross-dependencies built in parallel
  - Distcc/gmake -j<n> provides file level parallelism within package
    - Several distcc servers do compilation after macro/header expansion
  - Incremental builds
    - Only checkout changed packages and rebuilt them and clients

NICOS global page

http://atlas-computing.web.cern.ch/atlas-computing/links/distDirectory/nightlies/global/

ATLAS NICOS RTT FCT Tag BBC News Cars (2) Experiments Laboratories Mac (13) Misc (30) Online Accounts Technology (151)

NICOS | Draft agenda: ATL... | Tier-0 Monitoring ... | Atlas Generators -... | Swisscom - Reside... | http://atlastagcoll... | HepMC - a C++ Ev...

NICOS NIGHTLIES SUMMARY

NEW: build nodes info

Last modified Wed, 02 Jul 2008 06:05:07 GMT

This web page shows the list of ATLAS nightlies (see twiki page for details). Some nightlies are not added to this list: experimental nightlies, MAC nightlies, doxygen builds as well as stable releases.

Nightly testing summaries: ATN (rel_0, rel_1, rel_2, rel_3, rel_4, rel_5, rel_6, also accessible from nightly web pages), RTT, FCT (AID scans).

Message of a day about nightlies problems and usability

```
30 May 2008: power outage, most Friday's nightlies are not completed
24 - 27 May 2008: kernel upgrades caused occasional nightlies failures (in particular 14.1.0.Y, LCG3, MIG0)
20 May 2008: development validation rel_2 nightlies delayed by 6 hours (to get new AthenaServices version)
12 May 2008: development rel_6 and rel_0 nightlies failed, rel_1 delayed
because of looping during compilation of one of AtlasTrigger packages
```

Nightly Title	# Platforms	# Projects	Latest Rel.	Build	Date	Copy	Ave. Failed Builds	Ave. Test Success(%)
MAJOR NIGHTLIES								
14.2.X	4	11	rel_3	done	07/02 07:51	done	0.5	71.9
14.2.X-VAL	2	11	rel_3	done	07/02 07:22	done	0.5	76.6
Kt_14.2.X	2	1	rel_2	done	07/01 12:39	N/A	0	0
ATLASPOINT1 AND ATLASPRODUCTION NIGHTLIES								
13.0.40.Y	2	1	rel_2	done	07/01 06:40	done	0	30.5
14.0.10.Y	2	1	rel_3	done	07/02 04:04	done	0	57.0
14.1.0.Y	2	1	rel_3	done	07/02 05:58	done	0	0
14.1.0.Y-VAL	1	1	rel_3	done	07/02 03:07	done*	0	0
14.2.0.Y	1	1	rel_3	done	07/02 00:55	done	0	52.0
14.2.0.Y-VAL	1	1	rel_2	done	07/01 14:57	done	0	67.0
14.2.5.Y	2	1	rel_3	done	07/02 00:05	done	0	0
Kt_Pnt_14.0.10.Y	1	1	rel_3	done	07/02 04:19	N/A	0	0
Kt_Pnt_14.1.0.Y	1	1	rel_3	done	07/02 06:23	N/A	0	0
Kt_Pnt_14.1.0.Y-VAL	1	1	rel_3	done	07/02 05:00	N/A	1	0
Kt_Pnt_14.2.5.Y	1	1	rel_3	done	07/02 01:28	N/A	0	0
Kt_Prod_13.0.40.Y	2	1	rel_2	done	07/01 04:59	N/A	0	0
Kt_Prod_14.2.0.Y	1	1	rel_3	done	07/02 02:44	N/A	0	0
OTHER NIGHTLIES								
14.2.X-LCG1	1	10	rel_1	done	06/30 18:18	done	121.1	5
14.2.X-LCG3	1	10	rel_2	done	07/01 17:56	done	125.0	23.5
14.2.X-MIG0	1	11	rel_1	done	07/02 04:09	done	0.5	70.1
14.2.X-MIG1	1	10	rel_1	done	07/02 06:02	done	1	56.5

NICOS common page

http://atlas-computing.web.cern.ch/atlas-computing/links/distDirectory/nightlies/global/index14

ATLAS NICOS RTT FCT Tag BBC News Cars (2) Experiments Laboratories Mac (13) Misc (30) Online Accounts Technology (151)

NICOS common page | Draft agenda: ATL... | Tier-0 Monitoring ... | Atlas Generators -... | Swisscom - Reside... | http://atlastagcoll... | HepMC - a C++ Ev...

NICOS version 1_2_0

NICOS PROJECTS FOR 14.2.X-VAL NIGHTLIES

full list of tags

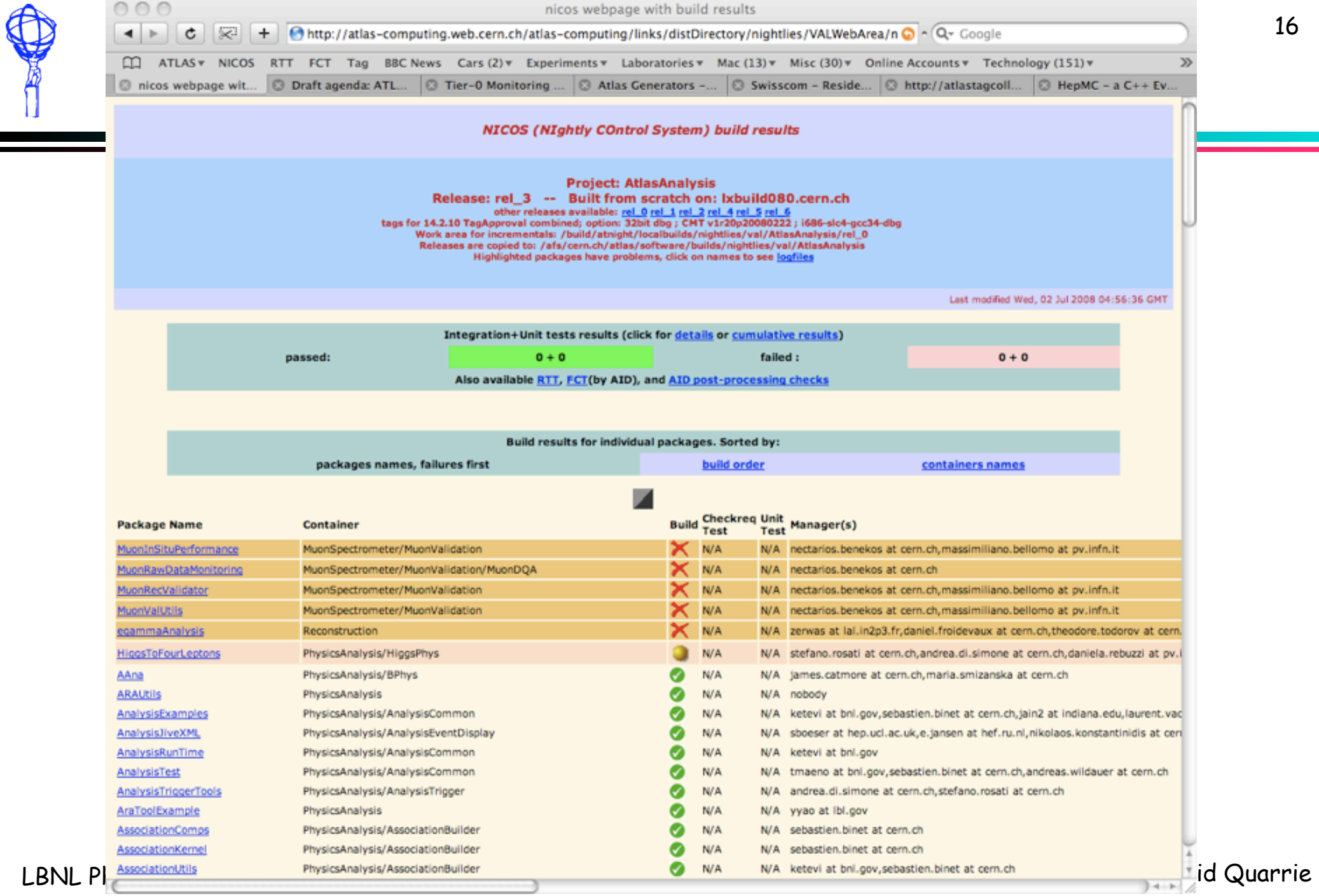
Last modified Wed, 02 Jul 2008 05:53:38 GMT

Project	Platform	Latest	Build	Copy	Test	Date	Failed Bids	Test OK(%)	NICOS Suffix
AtlasCore	i686-slc4-gcc34-dbg	rel_3	done	done	done	07/01 23:19	0	N/A	VAL142X32BS4CoreDbg
	i686-slc4-gcc34-opt	rel_3	done	done	done	07/02 00:26	0	100	VAL142X32BS4CoreOpt
DetCommon	i686-slc4-gcc34-dbg	rel_3	done	done	done	07/01 23:20	0	N/A	VAL142X32BS4DetDbg
	i686-slc4-gcc34-opt	rel_3	done	done	done	07/01 23:32	0	0	VAL142X32BS4DetOpt
AtlasConditions	i686-slc4-gcc34-dbg	rel_3	done	done	done	07/02 00:23	0	N/A	VAL142X32BS4CondDbg
	i686-slc4-gcc34-opt	rel_3	done	done	done	07/02 00:46	0	100	VAL142X32BS4CondOpt
AtlasEvent	i686-slc4-gcc34-dbg	rel_3	done	done	done	07/02 01:36	0	N/A	VAL142X32BS4EvtDbg
	i686-slc4-gcc34-opt	rel_3	done	done	done	07/02 02:11	0	81	VAL142X32BS4EvtOpt
AtlasReconstruction	i686-slc4-gcc34-dbg	rel_3	done	done	done	07/02 03:26	0	N/A	VAL142X32BS4RecDbg
	i686-slc4-gcc34-opt	rel_3	done	done	done	07/02 05:48	0	78	VAL142X32BS4RecOpt
AtlasSimulation	i686-slc4-gcc34-dbg	rel_3	done	done	done	07/02 04:08	0	N/A	VAL142X32BS4SimDbg
	i686-slc4-gcc34-opt	rel_3	done	done	done	07/02 07:36	0	85	VAL142X32BS4SimOpt
AtlasTrigger	i686-slc4-gcc34-dbg	rel_3	done	done	done	07/02 05:41	0	N/A	VAL142X32BS4TrgDbg
	i686-slc4-gcc34-opt	rel_3	done	done	done	07/02 07:35	0	45	VAL142X32BS4TrgOpt
AtlasAnalysis	i686-slc4-gcc34-dbg	rel_3	done	done	done	07/02 06:56	5	N/A	VAL142X32BS4AnlDbg
	i686-slc4-gcc34-opt	rel_3	done	done*	work	07/02 07:05	5	N/A	VAL142X32BS4AnlOpt
AtlasOffline	i686-slc4-gcc34-dbg	rel_3	done	done	done	07/02 07:07	0	N/A	VAL142X32BS4OfflDbg
	i686-slc4-gcc34-opt	rel_3	done	done*	work	07/02 07:16	0	N/A	VAL142X32BS4OfflOpt
AtlasProduction	i686-slc4-gcc34-dbg	rel_3	done	done	done	07/02 07:10	0	N/A	VAL142X32BS4ProdDbg
	i686-slc4-gcc34-opt	rel_3	done	done	done	07/02 07:20	0	100	VAL142X32BS4ProdOpt
AtlasPoint1	i686-slc4-gcc34-dbg	rel_3	done	done	done	07/02 07:11	0	N/A	VAL142X32BS4PntDbg
	i686-slc4-gcc34-opt	rel_3	done	done	done	07/02 07:22	0	100	VAL142X32BS4PntOpt

NEW: the "globe and envelope symbol" indicate that the automatic e-mails about build or test problems are sent

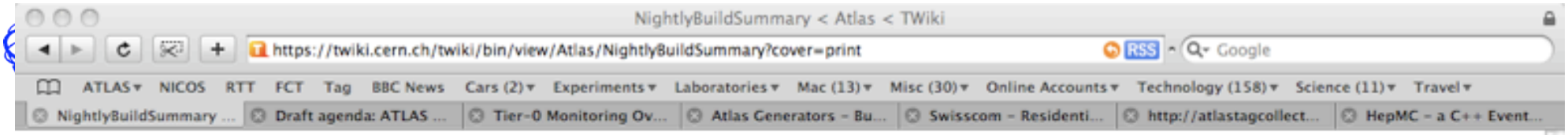
- **This page** shows the list of 14.2.X-VAL nightly projects managed by the NICOS system
- Click here to get a list of all nightlies available
- Click on the platform name to get the results for all releases available (for particular project and platform)
- Click on the latest release name to get the summary of build and test results for the upstream project (for particular platform)

Send comments or questions about NICOS


nicos webpage with build results
http://atlas-computing.web.cern.ch/atlas-computing/links/distDirectory/nightlies/VALWebArea/n
Google
ATLAS NICOS RTT FCT Tag BBC News Cars (2) Experiments Laboratories Mac (13) Misc (30) Online Accounts Technology (151)
nicos webpage wit... Draft agenda: ATL... Tier-0 Monitoring ... Atlas Generators -... Swisscom - Reside... http://atlastagcoll... HepMC - a C++ Ev...
NICOS (NIghtly COntrol System) build results
Project: AtlasAnalysis
Release: rel_3 -- Built from scratch on: lxbuild080.cern.ch
other releases available: rel_0 rel_1 rel_2 rel_4 rel_5 rel_6
tags for 14.2.10 TagApproval combined; option: 32bit dbg ; CMT v1r20p20080222 ; i686-slc4-gcc34-dbg
Work area for incrementals: /build/atnight/localbuilds/nightlies/val/AtlasAnalysis/rel_0
Releases are copied to: /afs/cern.ch/atlas/software/builds/nightlies/val/AtlasAnalysis
Highlighted packages have problems, click on names to see logfiles
Last modified Wed, 02 Jul 2008 04:56:36 GMT
Integration+Unit tests results (click for details or cumulative results)
passed: 0 + 0
failed : 0 + 0
Also available RTT, FCT(by AID), and AID post-processing checks
Build results for individual packages. Sorted by:
packages names, failures first
build order
containers names
Package Name
Container
Build
Checkreq Test
Unit Test
Manager(s)
MuonInSituPerformance MuonSpectrometer/MuonValidation N/A N/A nectarios.benekos at cern.ch,massimiliano.bellomo at pv.infn.it
MuonRawDataMonitoring MuonSpectrometer/MuonValidation/MuonDQA N/A N/A nectarios.benekos at cern.ch
MuonRecValidator MuonSpectrometer/MuonValidation N/A N/A nectarios.benekos at cern.ch,massimiliano.bellomo at pv.infn.it
MuonValUtils MuonSpectrometer/MuonValidation N/A N/A nectarios.benekos at cern.ch,massimiliano.bellomo at pv.infn.it
egammaAnalysis Reconstruction N/A N/A zerwas at lal.in2p3.fr,daniel.froidevaux at cern.ch,theodore.todorov at cern.
HiggsToFourLeptons PhysicsAnalysis/HiggsPhys N/A N/A stefano.rosati at cern.ch,andrea.di.simone at cern.ch,daniela.rebuzzi at pv.
AAna PhysicsAnalysis/BPhys N/A N/A james.catmore at cern.ch,maria.smizanska at cern.ch
ARAUtils PhysicsAnalysis N/A N/A nobody
AnalysisExamples PhysicsAnalysis/AnalysisCommon N/A N/A ketevi at bnl.gov,sebastien.binet at cern.ch,jain2 at indiana.edu,laurent.vac
AnalysisJiveXML PhysicsAnalysis/AnalysisEventDisplay N/A N/A sboeser at hep.ucl.ac.uk,e.jansen at hef.ru.nl,nikolaos.konstantinidis at cer
AnalysisRunTime PhysicsAnalysis/AnalysisCommon N/A N/A ketevi at bnl.gov
AnalysisTest PhysicsAnalysis/AnalysisCommon N/A N/A tmaeno at bnl.gov,sebastien.binet at cern.ch,andreas.wildauer at cern.ch
AnalysisTriggerTools PhysicsAnalysis/AnalysisTrigger N/A N/A andrea.di.simone at cern.ch,stefano.rosati at cern.ch
AraToolExample PhysicsAnalysis N/A N/A yyao at lbl.gov
AssociationComps PhysicsAnalysis/AssociationBuilder N/A N/A sebastien.binet at cern.ch
AssociationKernel PhysicsAnalysis/AssociationBuilder N/A N/A sebastien.binet at cern.ch
AssociationUtils PhysicsAnalysis/AssociationBuilder N/A N/A ketevi at bnl.gov,sebastien.binet at cern.ch

Special Nightlies

- Goal is to keep the primary nightlies as stable as possible
- Validation nightlies
  - Tags that have been put into validation (rather than accepted) run in a special validation nightly
    - E.g. 14.2.10-VAL
  - If there are no side effects and requests work as expected, they can be accepted
- Migration nightlies
  - Branches in Tag Collector where package tags override those in the primary nightlies
    - E.g. 14.2.X-MIG1
  - Used for disruptive migrations
    - E.g. For new GAUDI versions; EDM changes etc.
- LCG nightlies
  - Builds against LCG-AppsArea nightly builds (external packages)



Introduction

This page summarizes the nightly builds including the special purpose and migration nightly builds, in particular the role (which varies as a function of time), details (how many days the cycle covers, the CMTCONFIG version(s)) and the responsible person(s). The **TAG** column indicates the AtlasLogin environment setup tag that should be specified in order to access the corresponding nightly. The **Deadline** for tag submission is given in the CERN timezone. The primary NICOS page is located here.

Primary Nightly Builds

NAME	CYCLE	Deadline	CMTCONFIG	TAG	DESCRIPTION	COORDINATOR(S)
14.2.X	rel_0-rel_6	21:00	i686-slc4-gcc34-opt/dbg, x84_64-slc4-gcc34-opt/dbg	bugfix	Primary nightlies leading up to 14.2.10	EmilObreshkov, DmitryEmelyanov
14.2.X-VAL	rel_0-rel_6	21:00	i686-slc4-gcc34-opt/dbg	bugfix,val	Package tags undergoing validation	EmilObreshkov, DmitryEmelyanov

Patch Nightly Builds

NAME	CYCLE	Deadline	CMTCONFIG	TAG	DESCRIPTION	COORDINATOR(S)
14.2.0.Y	rel_0-rel_6	21:00	i686-slc4-gcc34-opt	14.2.0.Y	Patches for 14.2.0.Y (AtlasProduction)	AndresPacheco, ManuelGallas
14.2.0.Y-VAL	rel_0-rel_6	11:00	i686-slc4-gcc34-opt	14.2.0.Y-VAL	Validation for 14.2.0.Y (AtlasProduction)	AndresPacheco, ManuelGallas
14.1.0.Y	rel_0-rel_6	21:00	i686-slc4-gcc34-opt/dbg	14.1.0.Y	Patches for 14.1.0.50+ (AtlasPoint1) M7 Tier0 Re-processing	MariaCosta ?, JamieBoyd
14.1.0.Y-VAL	rel_0-rel_6	10:00	i686-slc4-gcc34-opt	14.1.0.Y-VAL	Patches/Validation for 14.1.0.14+ (AtlasPoint1) FDR2 Tier0 Processing	DavidRousseau, SebastianSchaetzel
14.1.0.Y	rel_0-rel_6	19:45	i686-slc4-gcc34-opt	14.1.0.Y	Patches for 14.1.0.Y (AtlasProduction)	AndresPacheco, ManuelGallas
13.0.40.Y	rel_0-rel_6	18:00	i686-slc3-gcc323-opt, i686-slc4-gcc34-opt/dbg	point1 or pcache	Patches for 13.0.40 (AtlasProduction)	AndresPacheco

Special and Migration Nightly Builds

The usage of these was last checked on 18 June 2008.

NAME	CYCLE	Deadline	CMTCONFIG	TAG	DESCRIPTION	COORDINATOR(S)
14.2.X-LCG	rel_0-rel_6	10:00	i686-slc4-gcc34-dbg, i686-slc4-gcc41-opt	lcg	Built against LCG "dev" slot (GAUDI ATLAS and LCG-preview)	EmilObreshkov, DavidQuarrie
14.2.X-LCG1	rel_0-rel_6	10:00	i686-slc4-gcc34-opt	lcg1	Built against LCG "dev1" slot (GAUDI ATLAS and LCG-55 - no SEAL)	EmilObreshkov, DavidQuarrie
14.2.X-LCG3	rel_0-rel_6	10:00	i686-slc4-gcc34-dbg	lcg3	Build against LCG "dev3" slot (GAUDI ATLAS and LCG_54-patches)	EmilObreshkov, DavidQuarrie
14.2.X-MIG0	rel_0-rel_1	21:00	i686-slc4-gcc34-dbg	mig0	Geant4.9.1 migration	AndreaDellAcqua, ZacharyMarshall
14.2.X-MIG1	rel_0-rel_1	21:00	i686-slc4-gcc34-dbg	mig1	Physics Analysis Development	SvenMenke, StathesPaganis
14.2.X-MIG2	rel_0-rel_1	21:00	i686-slc4-gcc34-opt	mig2	Tracking migrations (several)	AndreasWildauer
14.2.X-MIG3	rel_0-rel_1	21:00	i686-slc4-gcc34-opt	mig3	Calo EDM and Muon Combined Performance EDM migrations	KeteviAssamagan, GuillaumeUnal
14.2.X-MIG4	rel_0-rel_1	21:00	i686-slc4-gcc34-opt	mig4	AthenaBarCode Migration	YushuYao, PaoloCalafiura
14.2.X-MIG5	rel_0-rel_1	21:00	i686-slc4-gcc34-dbg	mig5	TileConditions migrations (Free from now on 23 June)	Main.SashaSolodkov
14.2.X-MIG6	rel_0-rel_1	21:00	i686-slc4-gcc34-dbg	mig6	Trigger EDM Migration (Free from now on 23 June)	JohnBaines

Instructions for Developers submitting tags to Special and Migration Nightlies

The normal Tag Collector tag approval procedures can be used in order to allow developers to submit tags to these special nightlies. The instructions are identical to those for inserting tags into patch projects.

Instructions for Release Coordinators

For each of the migration and special purpose nightlies listed above, someone will be designated to be the Release Coordinator. This will typically be for a finite duration, being the time taken by a single migration. The release coordination has control over adding, updating and removing packages in these branches. Any package and version that is inserted into such a branch will override the equivalent package and version in the baseline branch. Such

Validation/Test Scaffolds

- ATN (ATNight)
  - Unit and component tests run inline with release builds
    - Restricted statistics since run on single machine
- RTT (RunTimeTester)
  - ~20 Linux machines dedicated to tests (~100s of tests)
  - Typically ~few hours
- FCT (Full Chain Tests)
  - Tests that cover all production jobs (~10s of tests)
    - Generation, simulation, digitization, reconstruction, etc.
  - Scale of ~100s events (~4 hours)
- Tier0 Chain Tests (TCT)
  - Tests that replicate Tier0 processing
    - Reconstruction, monitoring, merging, TAG and DPD creation etc.
    - Scale of ~10,000-100,000 events

RTT runs for the past week

http://atlas-project-rtt-results.web.cern.ch/atlas-project-rtt-results/weeklyTable/weekly.php

ATLAS NICOS RTT FCT Tag BBC News Cars (2) Experiments Laboratories Mac (13) Misc (30) Online Accounts Technology (159) Science (11) Travel

RTT | Draft agenda: ATLAS ... | Tier-0 Monitoring Ov... | Atlas Generators - Bu... | Swisscom - Residenti... | http://atlastagcollect... | HepMC - a C++ Event...

RTT performance over the past week

[Last update (CERN time): 09:20 @ 02 Jul 2008]

Show All Reports | Show All Results

Release	Branch	Platform	Heirarchy	Started	Completed	Jobs OK	Done	Total	Pkgs: jobs OK	Report
Wednesday 02 July 2008										
rel_3	14.0.10.Y	i686-slc4-gcc34-opt	AtlasPoint1	08/07/02@03:00	08/07/02@05:33	3	4	4	59/60	---
---	14.1.0.Y	i686-slc4-gcc34-opt	AtlasProduction	---	---	---	---	---	---/---	---
---	bugfix	i686-slc4-gcc34-opt	AtlasHLT	---	---	---	---	---	---/---	---
rel_3	14.2.0.Y	i686-slc4-gcc34-opt	AtlasProduction	08/07/02@00:15	n/a	225	351	513	28/68	---
---	bugfix	i686-slc4-gcc34-opt	AtlasProduction	---	---	---	---	---	---/---	---
---	val	i686-slc4-gcc34-opt	AtlasProduction	---	---	---	---	---	---/---	---
rel_3	14.1.0.Y	i686-slc4-gcc34-opt	AtlasPoint1	08/07/02@00:35	08/07/02@05:46	1	3	3	61/62	---
Tuesday 01 July 2008										
rel_2	14.0.10.Y	i686-slc4-gcc34-opt	AtlasPoint1	08/07/01@03:00	08/07/01@05:52	2	4	4	59/60	---
---	14.1.0.Y	i686-slc4-gcc34-opt	AtlasProduction	---	---	---	---	---	---/---	---
rel_2	bugfix	i686-slc4-gcc34-opt	AtlasHLT	08/07/01@07:30	08/07/01@11:09	8	18	18	30/32	---
rel_2	14.2.0.Y	i686-slc4-gcc34-opt	AtlasProduction	08/07/01@00:15	08/07/01@16:46	431	513	513	32/68	---
rel_2	bugfix	i686-slc4-gcc34-opt	AtlasProduction	08/07/01@05:00	08/07/01@21:32	401	527	527	26/68	---
rel_2	val	i686-slc4-gcc34-opt	AtlasProduction	08/07/01@09:01	08/07/01@20:14	147	527	527	14/68	---
rel_2	14.1.0.Y	i686-slc4-gcc34-opt	AtlasPoint1	08/07/01@00:35	08/07/01@06:08	0	3	3	61/62	---
Monday 30 June 2008										
rel_1	14.0.10.Y	i686-slc4-gcc34-opt	AtlasPoint1	08/06/30@03:00	08/06/30@03:36	2	4	4	59/60	Show Report
---	14.1.0.Y	i686-slc4-gcc34-opt	AtlasProduction	---	---	---	---	---	---/---	---
rel_1	bugfix	i686-slc4-gcc34-opt	AtlasHLT	08/06/30@07:30	08/06/30@10:53	0	18	18	30/32	Show Report
rel_1	14.2.0.Y	i686-slc4-gcc34-opt	AtlasProduction	08/06/30@00:15	08/06/30@16:16	427	513	513	30/68	Show Report

Tier0 Bug Tracking

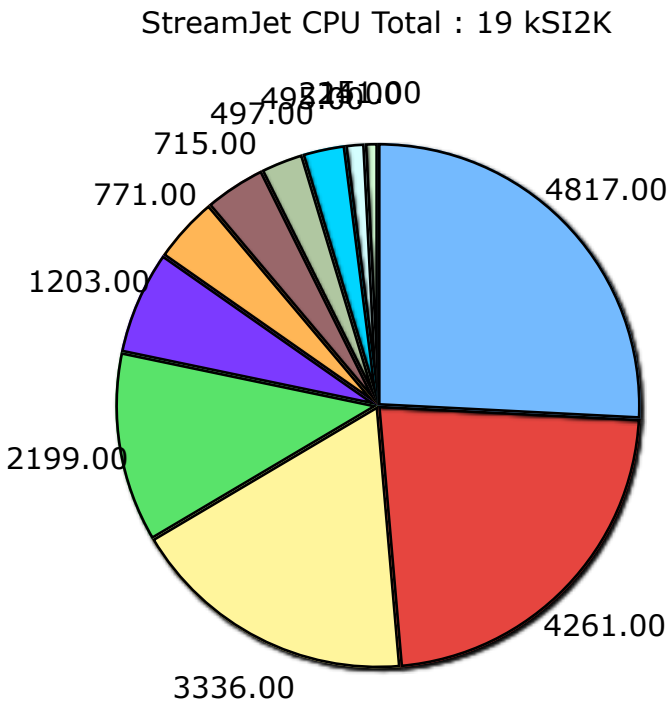
- Deal with initial turn-on problems
  - Unique period where software exposed to real physics events for first time
  - Hopefully won't need such a heavyweight system for subsequent running periods
- Dedicated Bug Tracking System (based on Savannah)
  - Bug reports submitted by Tier0 shifters
  - Daily shift crew of experienced validators
    - Remove duplicate bugs
    - Assign to appropriate domain (e.g. Inner Detector, Reconstruction)
    - Check on progress with open bugs (and apply pressure)
      - Backup of on-call experts and "SWAT" team
    - Submit tags corresponding to closed bugs to AtlasTier0 patch project
    - Check validation results
- 12-hour cycle of builds (daily & nightly)
  - Running Tier0 tests
- Daily sign-off meeting to decide on deployment of new patches

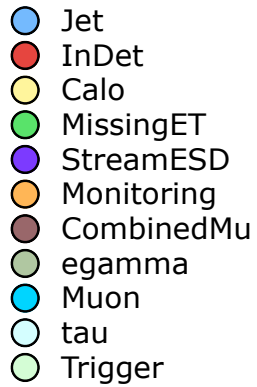
Reconstruction Performance and Descoping

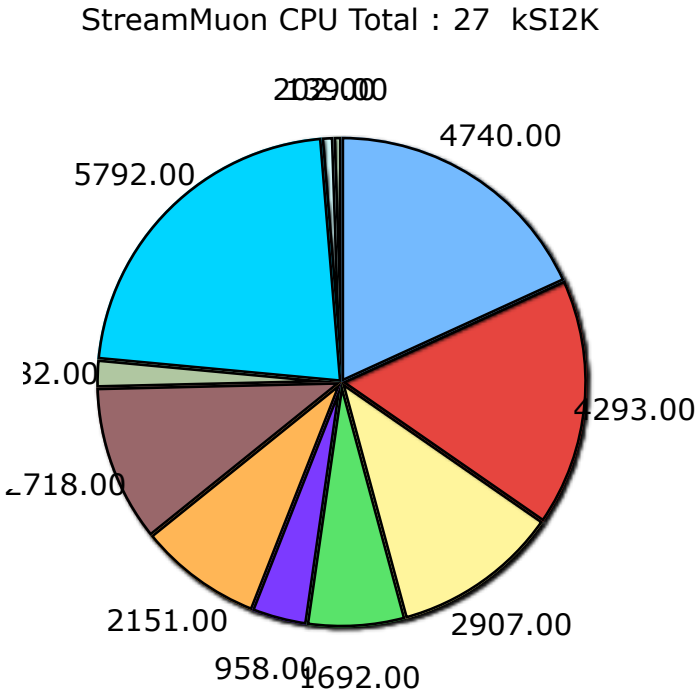
- Tier0 requirements are to handle 200Hz from ATLAS
- Tier0 sized based on 15kSi2K CPU secs per event (~6-7 seconds per typical core)
- Tier0 nodes have 2GB memory per core
  - No per-event swapping
- Crash rate should be as low as possible
  - <<1% of jobs
  - Each job ~10k events
    - <<1 crash per 10^6 events
  - There are about 100 reconstruction algorithms
    - <<1 crash per 10^8 algorithm-events

FDR-1 CPU Time Breakdown

- No pile up
  - but muon is more sensitive to luminosity than number of min bias per beam crossing⟹ should not be problem for 75 ns crossing this year



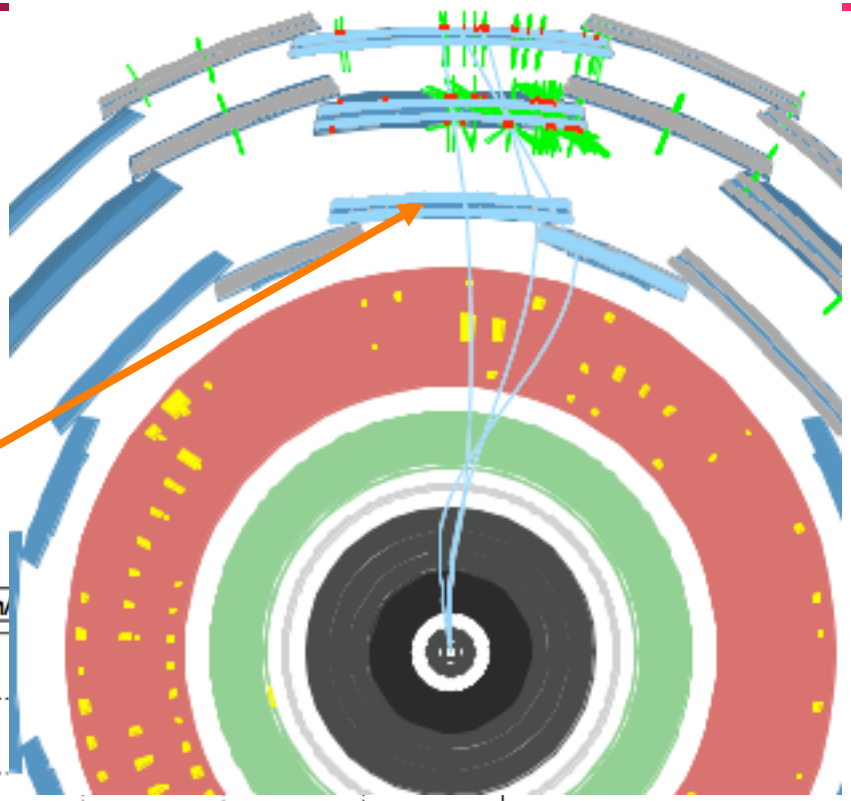


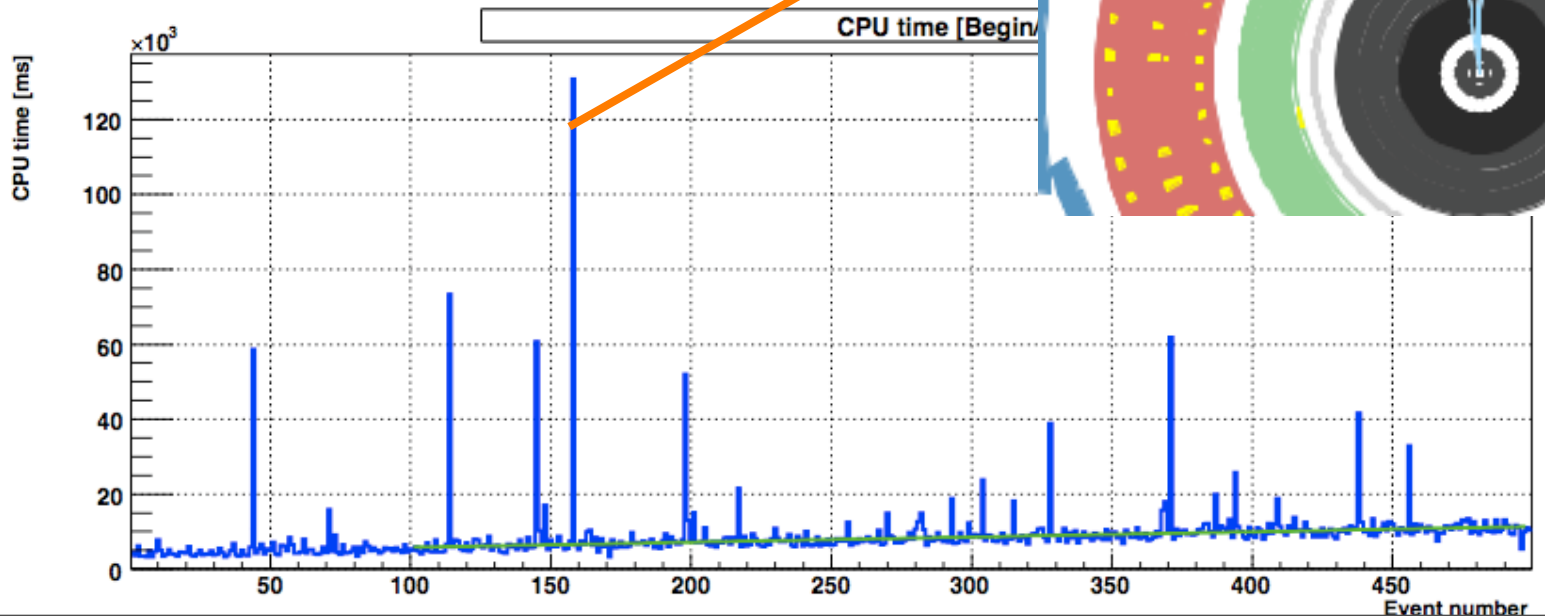


+AOD ~ 3kSi2K

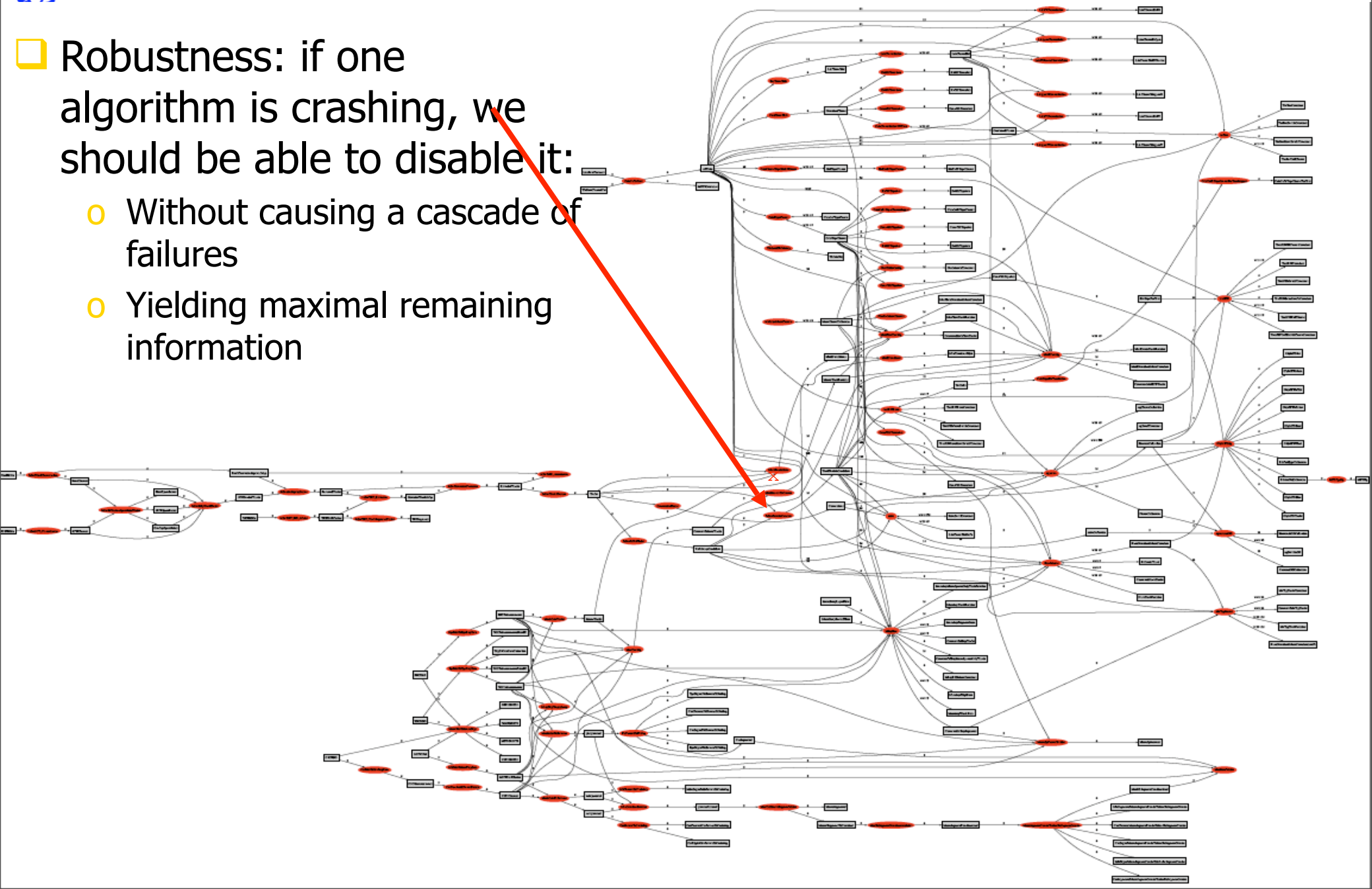
Commissioning

- M6 reco with 13.2.0 : ~30kSi2K s
- Fewer alg running than for standard atlas
- …but more difficult
  - Asynchronous
  - No vertex constraint
- CPU time per event spiky plot
  - (fixed since)





- Robustness: if one algorithm is crashing, we should be able to disable it:
  - Without causing a cascade of failures
  - Yielding maximal remaining information



Descoping handles

- Simplify one algorithm using existing properties
  - e.g. not run its more sophisticated tools
- Disable algorithm one by one or by logical grouping
- Disable persistency of a particular object
- Disable full sub-detector
- Fall back to a different condition DB tag
- Limit CPU time used per algorithm (if feasible)
- Disable reconstruction of part of a sub-detector (should be done at the RIO building stage)
- Split reconstruction in smaller jobs (i.e. spanning less than a luminosity block) (but this would be a last resort measure as this creates lots of complication downstream)
- One could also give up on the reconstruction of one stream (typically the jet one) (or not reconstruct all its luminosity blocks)

Example Failure Modes and Responses

- Jobs are crashing because of infrastructure problems (afs, castor, db server...):
  - No relevant descoping
- Jobs are crashing because the memory usage is well above 2GB already at the limit already at the beginning:
  - Disable/simplify algorithm according to priority list, and based on the algorithm memory consumption
- Jobs are crashing because the memory usage is well above 2GB at the end of the job because of a memory leak:
  - Disable the leaking algorithm(s), and if they could not be identified, disable/simplify algorithm according to priority list, and based on the algorithm memory consumption
  - ...or split reconstruction in smaller jobs
- Jobs are crashing in one monitoring task
  - Disable it this should be very very temporary
  - ...or split reconstruction in smaller jobs
- Jobs are crashing when persistifying a particular object
  - Disable the algorithm producing it or disable the persistency of the specific object

Descoping Priority List

- All detector and combined reconstruction groups asked to define their priority lists
  - Inner Detector
  - Muon and combined muons
  - Calorimetry
  - Jets
  - Missing-Et
  - Tau
  - E/gamma
  - B-tagging
- Persistency priority list (if events too large)
  - ESD can be reduced at expense of more access to RAW data
  - AOD can be reduced at expense of more frequent access to ESD

Descoping Gains and Issues

- Preliminary estimates for running only "high priority" algorithms, one saves
  - 400MB memory
  - 30% cpu time
  - 30% ESD size
  - 20% AOD size
- A further "essential" category could save x2 in cpu and disk size
- Need to further refine the priority list
- Need to ensure that all configuration handles to disable algorithms in place
  - At that downstream algorithms behave correctly
- Documented procedures for expert shifters to descope in the case of robustness problems (crashes) until a validated patch can be deployed
  - Crucial that Tier-0 continues to collect and distribute data
  - Tier-0 buffer can accommodate ~5 days worth of data until it fills up and introduces deadtime if data not already distributed

Stress Tests

- ATLAS has run stress tests of the software and computing infrastructure for many years
  - Data Challenges using simulated data
  - Approximately one per year
- New stress tests
  - Combined cosmics runs
    - Using the TDAQ and Tier-0 infrastructure
    - Typically every month for ~1 week
    - Not described in this talk, but important first exposure to real data from actual hardware
  - Full Dress Rehearsal (FDR)
    - Large scale tests using simulated data exercising data transfer from TDAQ through to physics analysis at Tier-2 centers and reprocessing at Tier-1s
  - Combined Computing Readiness Challenge
    - Test of primarily data distribution infrastructure with all LHC experiments operating simultaneously

Full Dress Rehearsal

- Series (2) of large scale tests using simulated data
- Mix events in realistic streams in bytestream format
  - Bulk physics, express and calibration streams
  - Introduce deficiencies
- Put data sample into output buffer disks of TDAQ system
  - Exercise data transfer protocol from TDAQ to Tier0
- Run calibration and data quality procedures on calibration/express streams
  - Exercise calibration loop upstream of first pass processing of bulk physics stream
    - Target latency of 24-48 hours
- Reconstruct and distribute persistent EDM
  - Event Summary Data (~1MB/event), Analysis Object Data (~200kB/event), Derived Physics Data (~10kB/event), TAG (~1kB/event)
- Physics Analysis at remote sites
- Bulk reprocessing at remote sites (improved calibrations and algorithms)

FDR Schedule

- FDR-1
  - Week of 4 February
  - 10 hours of 10^{31} data (+ 1 hour of 10^{32}): ~0.4 pb^{-1} of data
  - Release 12 simulated data, misaligned geometry
- FDR-2
  - Week of 2 June
  - A few hours of 10^{32} data (+ few minutes of 10^{33}): 1.5 pb^{-1} of data
  - Release 13 simulated data, mix of ideal and misaligned geometries
- Long gap between FDR-1 and FDR-2 allowed
  - New samples to be prepared
  - Incorporate the lessons learned
  - Lots of incremental improvements

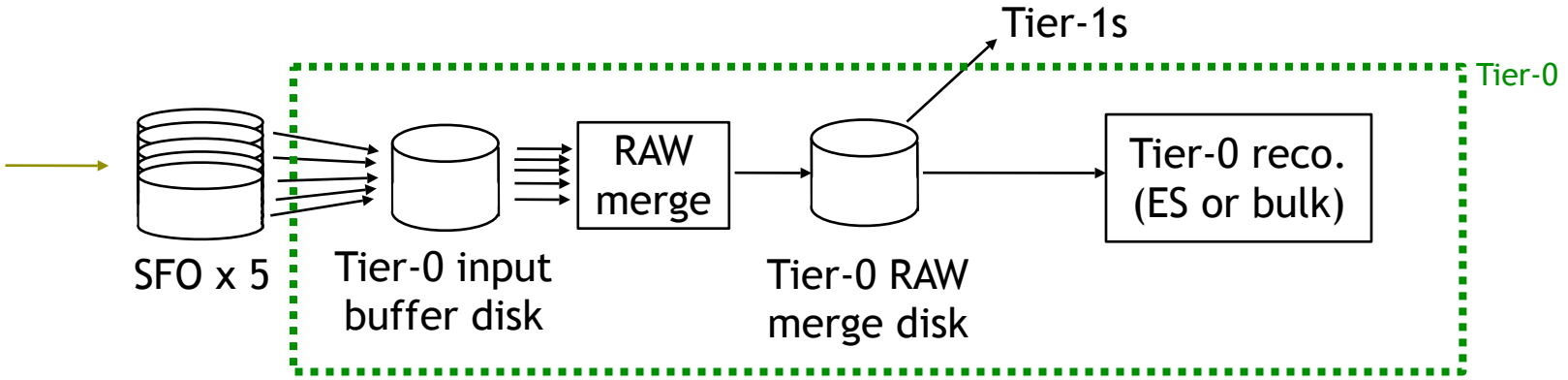
Data Samples

- Physics samples include
  - Minimum bias, single and doubly diffractive events
  - Random beam crossings
  - QCD jet events, photon-jet events
    - Including "biassed e-gamma fakes"
  - Drell-Yan, W and Z
  - b physics samples
  - Ttbar and single top
  - diboson
  - etc.

Data Streaming

- Implemented according to recommendations of Streaming Study Group
  - Streams persist through data processing to AOD creation
- Physics streams defined by High Level Trigger (HLT) bits
  - Muon + b-physics
  - e/gamma
  - Jets/tau/E_t^{miss}
  - Minbias
- Express stream
- Calibration streams
  - Inner detector alignment
    - High-p_t tracks, partial readout
    - Muon stream:Level 1 Muon Region Of Interest (ROI) partial readout)
- A "between-fill cosmics" sample for Inner Detector alignment

Raw Data Handling



SFOs (point 1)

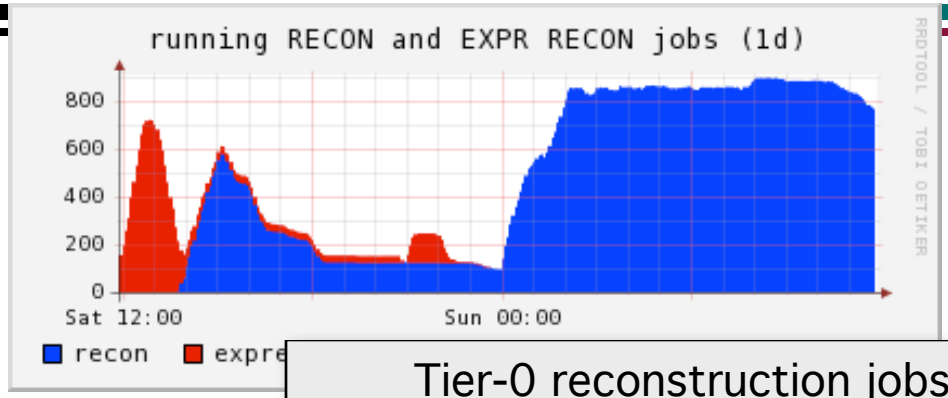
For FDR – replay data from SFOs

- files "appear" at SFO as though data
- one file per SFO, per stream, per lumi block

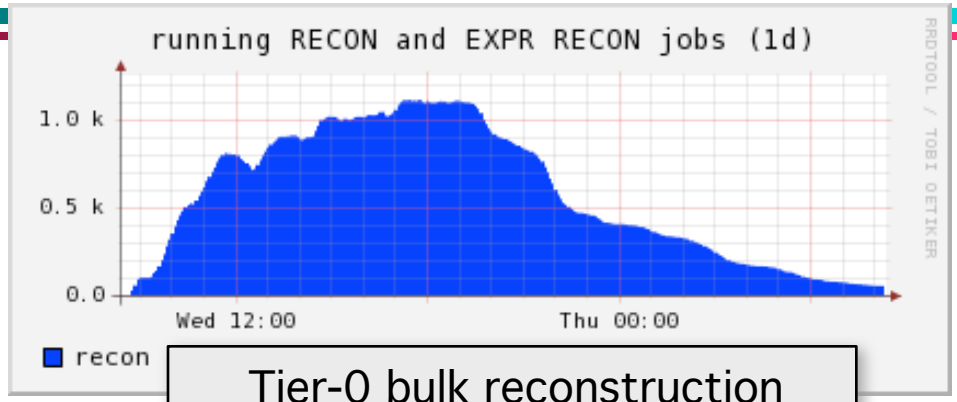
Metadata database (not shown) passes file-level info from SFOs to Tier-0 (FDR-2)

"RAW merge"joins BS files from 5 SFOs (same stream & lumi block) before export to Tier-1s (essential for Robustness of luminosity measurement)

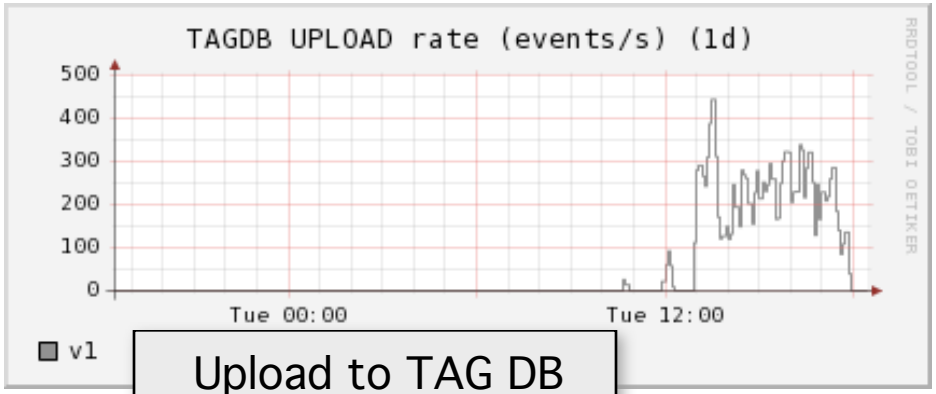
FDR-2 Process Monitoring



Tier-0 reconstruction jobs running over 24h period (Jun 7-8)

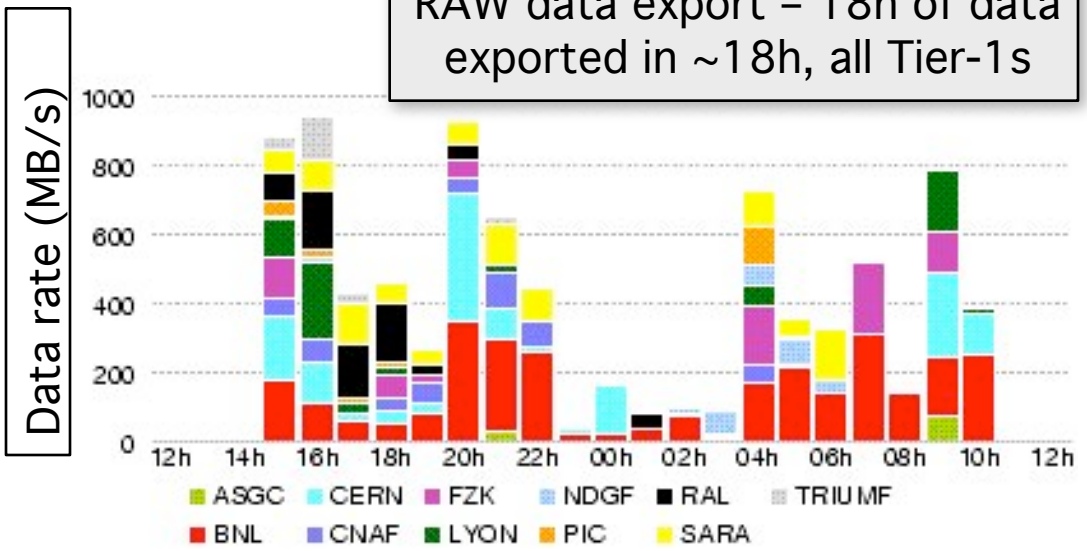


Tier-0 bulk reconstruction (Jun 11-12)



Upload to TAG DB (Jun 17)

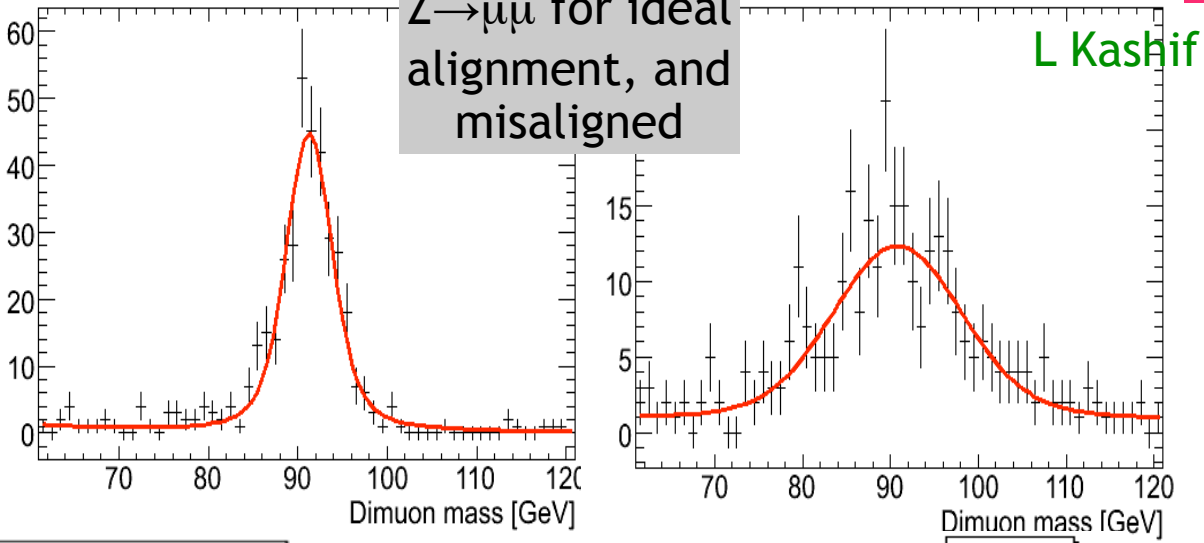
RAW data export – 18h of data exported in ~18h, all Tier-1s



FDR Analysis Examples

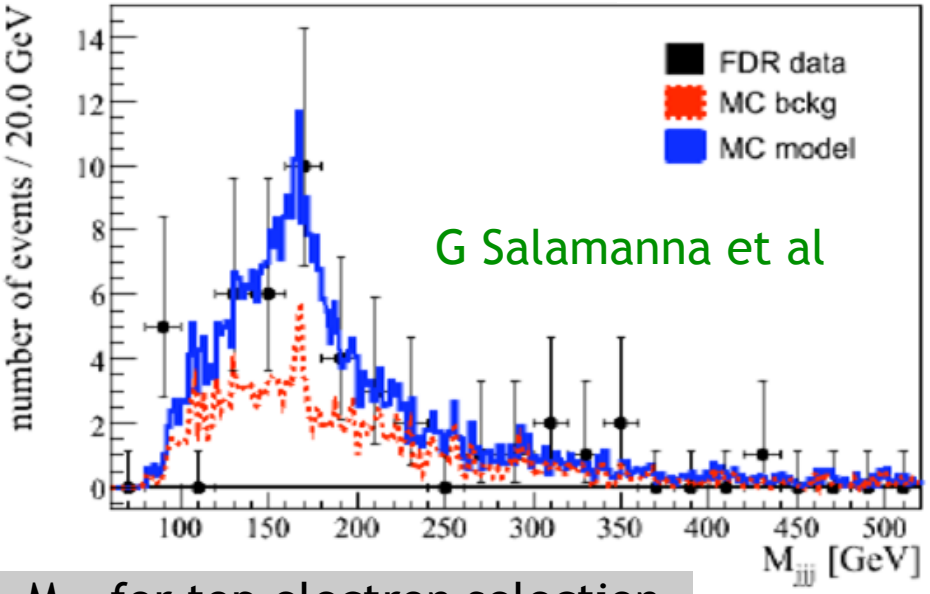
Some examples from recent FDR users meeting

Some cross-section bugs in FDR mixing – good exercise to track them down...?



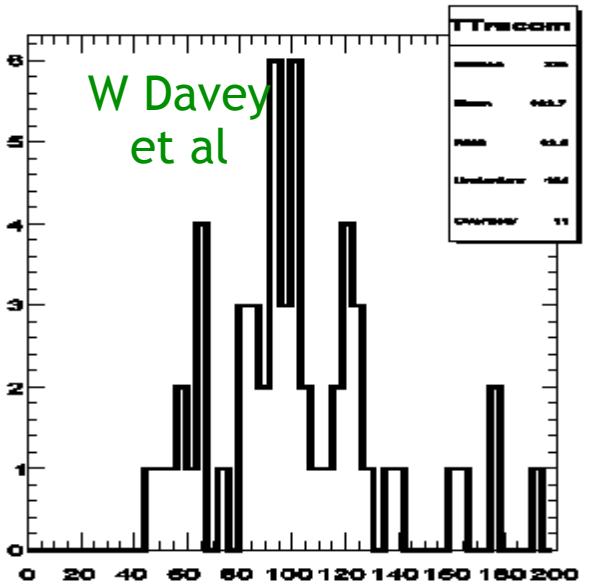
$Z \rightarrow \mu\mu$ for ideal alignment, and misaligned

L Kashif



G Salamanna et al

M_{jjj} for top electron selection



W Davey et al

$M(\tau\tau)$ candidates

FDR-2 Feedback

- Tier0 infrastructure in place
  - But work needed on automation and operation-by-shifters
- Basic Data Quality Monitoring infrastructure in place
  - But most detector subsystems need work to better define essential histograms etc.
- Tier-0 bug tracking system worked
  - But was heavily used because of lack of prior validation (6 patch releases within 1 week)
  - 5 bugs were found and fixed
- Reconstruction too slow but meets memory requirements
  - Need about 30% improvement in cpu time per event to meet 200Hz goal
    - Processing stages downstream of main reconstruction account for 30% of total
- Data export worked well
- More validation needed both prior to and after deployment
  - New tests being implemented now

Combined Computing Readiness Challenge

- CCRC'08 designed to test as many computing services of the experiment as possible
- All 4 experiments operating simultaneously
- CCRC08 Phase I:
  - Mostly a test of Storage Request manager (SRM) installation and configuration
  - Feb 2008 (some overlap with FDR-1)
- CCRC08 Phase II:
  - Full month of tests
    - No overlap with FDR
    - But some overlap with cosmics tests
    - CCRC08 only during week days
  - Focussed on data distribution
    - Rake payload data

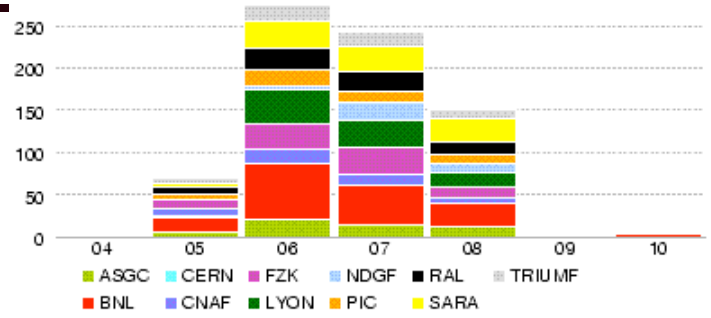
CCRC Tests

- Week 1: Data Distribution Functional Test
  - To make sure all files get where we want them to go
  - Between Tier0 and Tier-1s, for disk and tape
- Week 2: Tier-1 to Tier-1 tests
  - Similar rates as between Tier-0 and Tier-1
  - More difficult to control and monitor centrally
- Week 3: Throughput test
  - Try to maximize throughput but still following the model
  - Tier-0 to Tier-1 and Tier-1 to Tier-2
- Week 4: Finale, all tests together
  - Also artificial extra load from simulation production

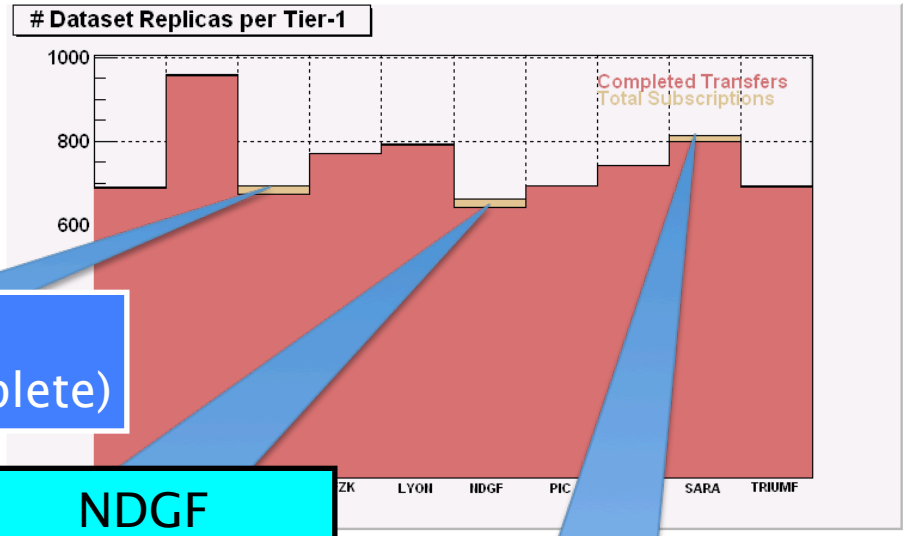
Week 1 - DDM Functional Tests

- Run load generator for 3 days at 40% of nominal rate
- Dataset subscribed to Tier-1 DISK and TAPE endpoints
  - RAW data subscribed according to ATLAS MoU shares (TAPE)
  - ESD subscribed only at the site hosting the parent RAW datasets (DISK)
    - In preparation for the Tier-1 to Tie-1 test of Week 2
  - AOD subscribed to every site (DISK)
- No activity for Tier-2s in Week 1
- Metrics:
  - Sites should hold a complete replica of 90% of subscribed datasets

Week-1 Results



CCRC08. May 2008 (week 19)



CNAF
(97% complete)

Temporary failure (disk server) treated as permanent by DDM. Transfer not retried

NDGF
(94% complete)

SARA
97% complete

Problematic throughput to TAPE

Limited resources:
1 disk buffer in front of 1 tape drive. Only 4 active transfers allowed. Clashes with FTS configuration (20 transfers).

Competition with other ("production") transfers

"Double Registration" problem

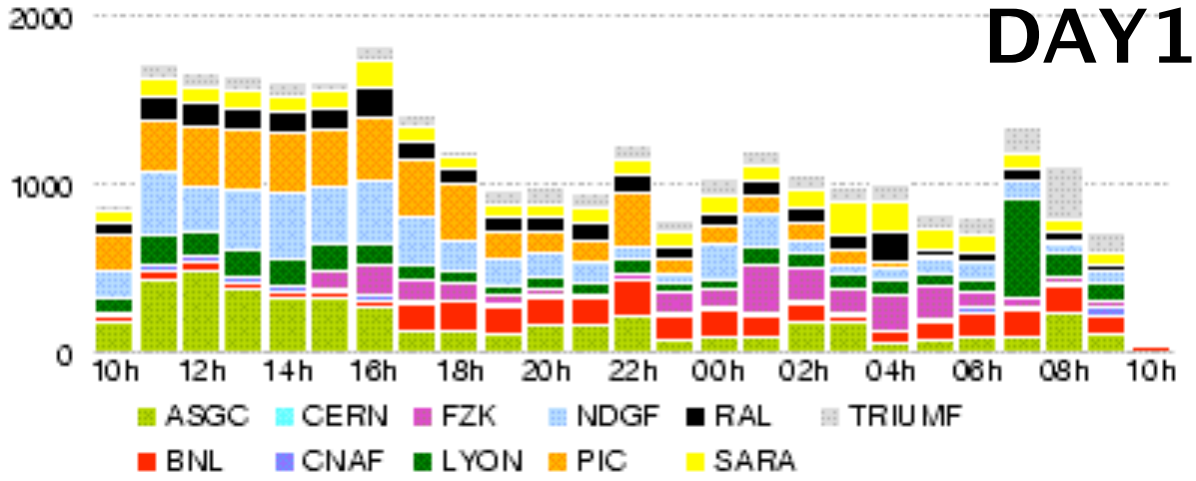
Slow transfer time out (changed from 600 to 3000s)
Storage fails to cleanup the disk pool after the entry of the file was removed from the namespace … disk full

Week 2: Tier-1 to Tier-1 Test

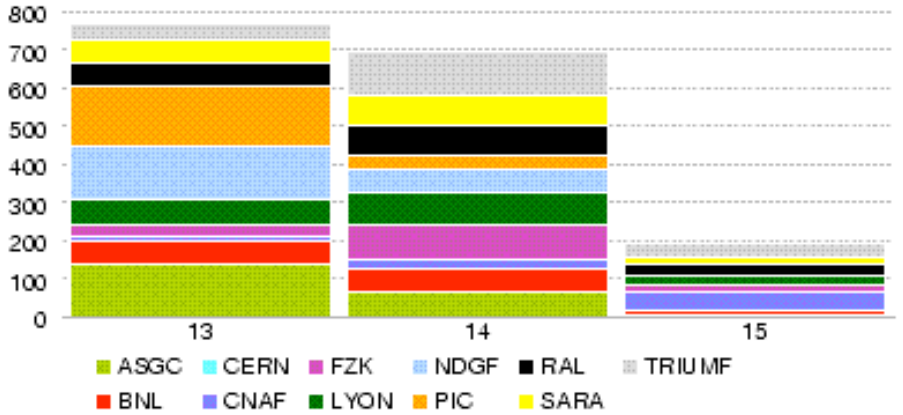
- Replicate ESD of Week 1 from "hosting Tier-1" to all other Tier-1s
  - Test of the full Tier-1 to Tier-1 transfer matrix
  - FTS at destination site schedules the transfer
  - Source site is always specified/imposed
    - No chaotic Tier-1 to Tie-1 replication - not in the ATLAS model
- Concurrent Tier-1 to Tier-1 exercise from CMS
  - Agreed in advance
- Dataset sample to be replicated
  - 629 datasets corresponding to 18TB of data
- Timing and Metrics
  - Subscriptions to every Tier-1 at 10:00 on May 13
    - All in one go - will the system throttle or collapse?
  - Exercise finishes at 14:00 on May 15
  - For every "channel" (T1-T1 pair) 90% of datasets should be completely transferred

Week-2: Results

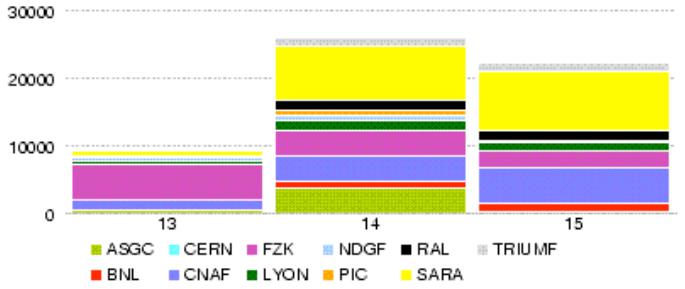
DAY1



All days (throughput)



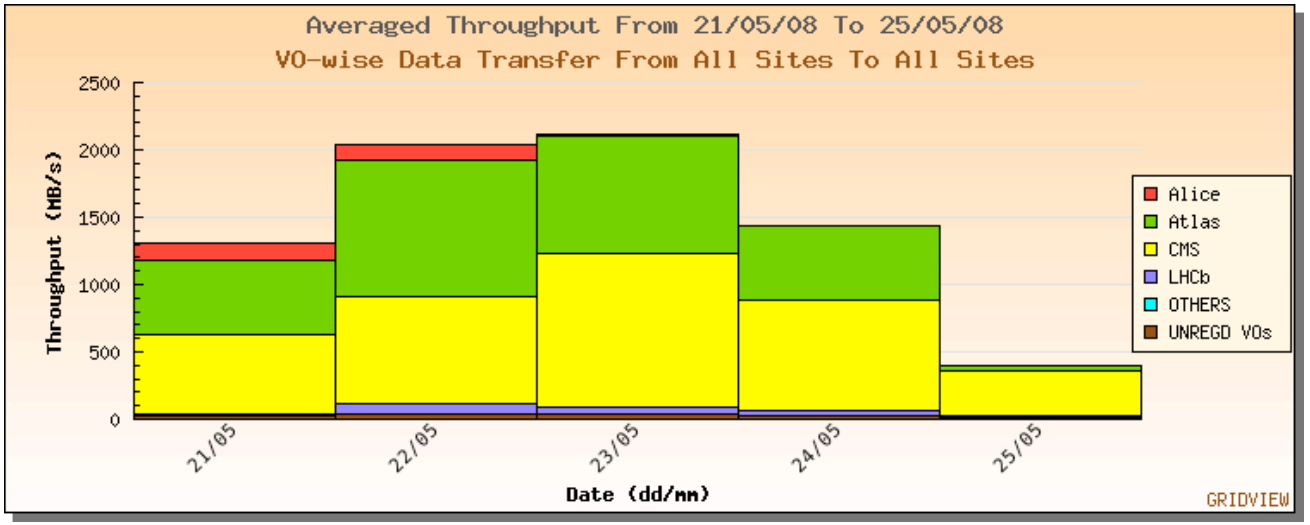
All days (errors)

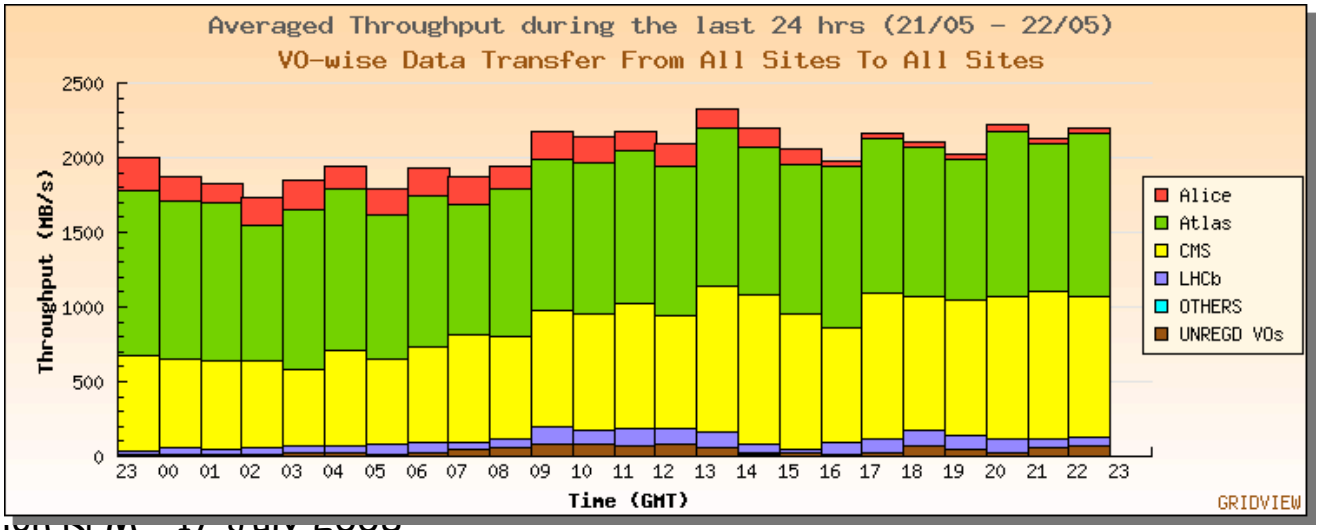


Week 3: Throughput Test

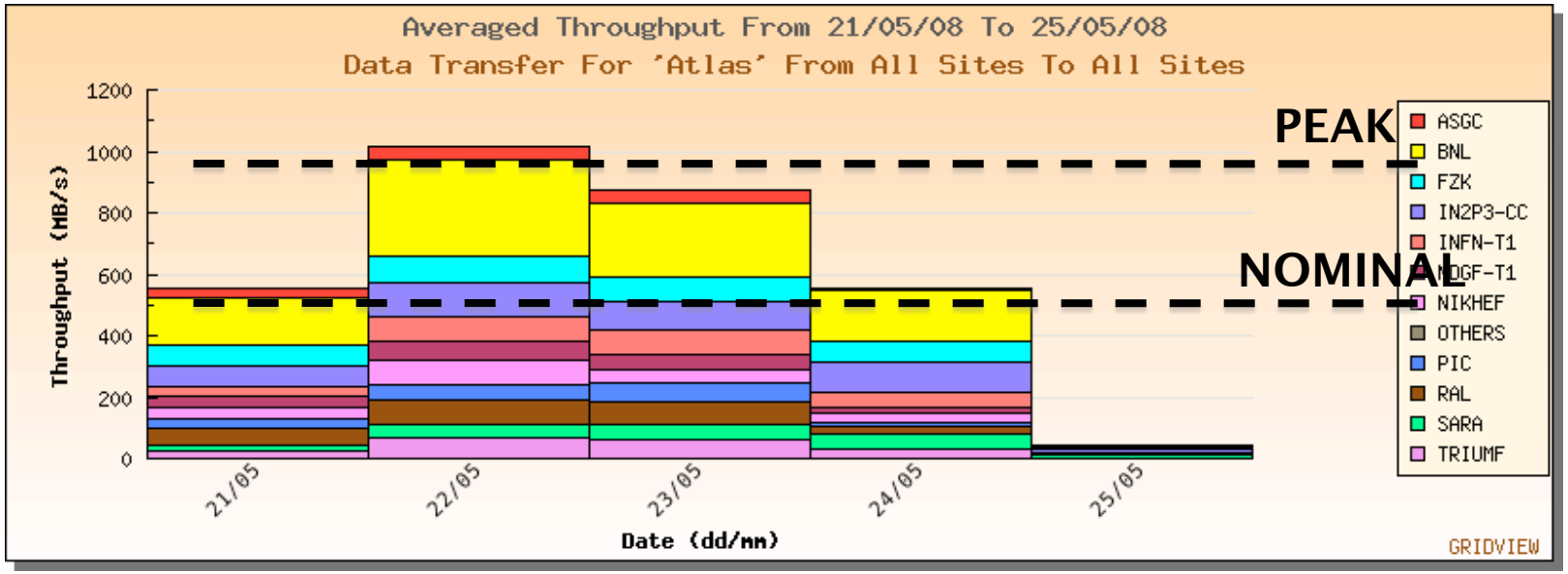
- Simulate data exports from Tier-0 for 24h/day of detector data taking at 200Hz
  - Nominal rate is 14h/day
- No over-subscription
  - Everything distributed according to computing model
- Timing and Metrics
  - Exercise starts at 10:0 on May 21 and ends at 10:00 on May 24
  - Sites should be able to sustain the peak rate for at least 24 hours and the nominal rate for 3 days

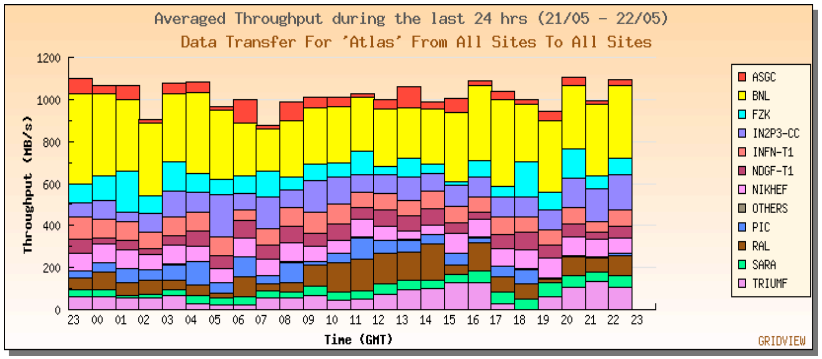
Week-3: all experiments

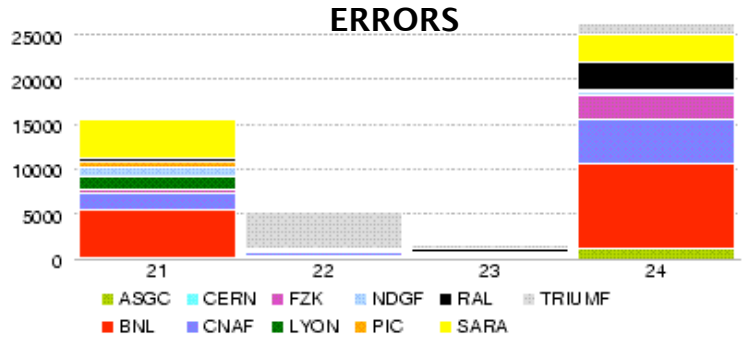




Week-3: Results







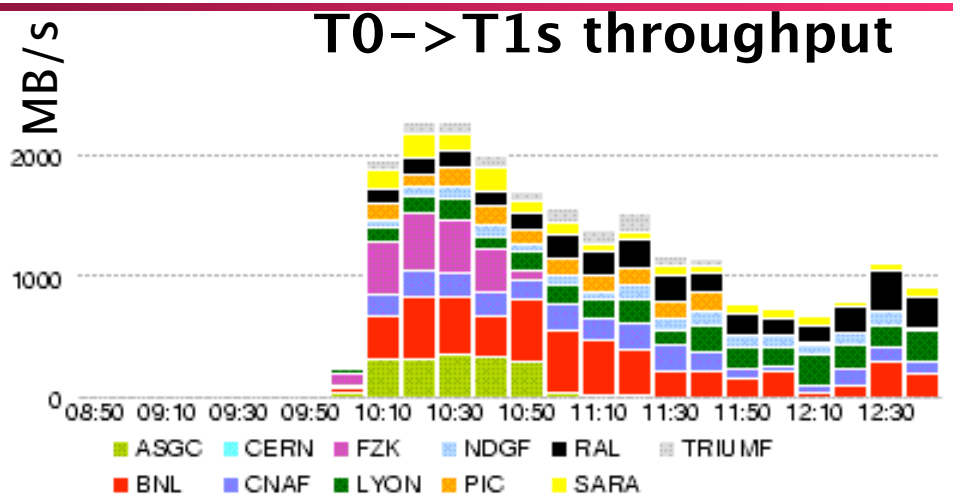
Week 4: Full Exercise

- The aim to test the full transfer matrix
  - Emulate the full load T0->T1 + T1->T1 + T1->T2
  - Considering the 14h data taking per dat
  - Considering the full stream reprocessing at 200Hz
- On top of this, add the burden of Monte Carlo production
  - Attempt to run a many jobs as one can
  - This also means transfers T1->T2 and T2-T1
- Four days exercise divided into two phases
  - First two days: functionality (lower rate)
  - Last two days: throughput (full steam)

Transfer ramp-up

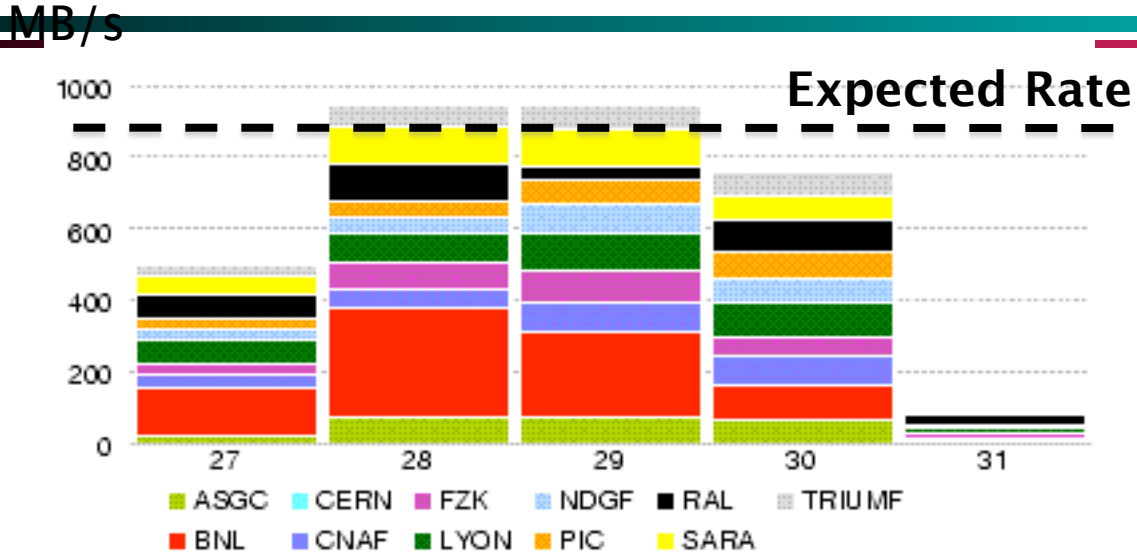
Test of backlog recovery
First data generated over 12 hours and subscribed in bulk

12h backlog recovered in 90 minutes!



	Transfers			Registrations		Errors	
Cloud	Efficiency	Throughput	Successes	Datasets	Files	Transfer	Registration
ASGC	100%	219 MB/s	300	46	300	0	0
BNL	100%	471 MB/s	597	10	597	0	0
CERN	0%	0 MB/s	0	0	0	0	0
CNAF	100%	195 MB/s	196	17	196	0	0
FZK	100%	229 MB/s	331	40	329	0	0
LYON	99%	147 MB/s	155	9	156	2	0
NDGF	100%	83 MB/s	98	22	98	0	0
PIC	100%	132 MB/s	156	19	156	0	0
RAL	99%	154 MB/s	152	17	152	1	0
SARA	100%	132 MB/s	207	16	208	0	0
TRIUMF	100%	105 MB/s	94	26	92	0	0

Throughputs



T0->T1 transfers

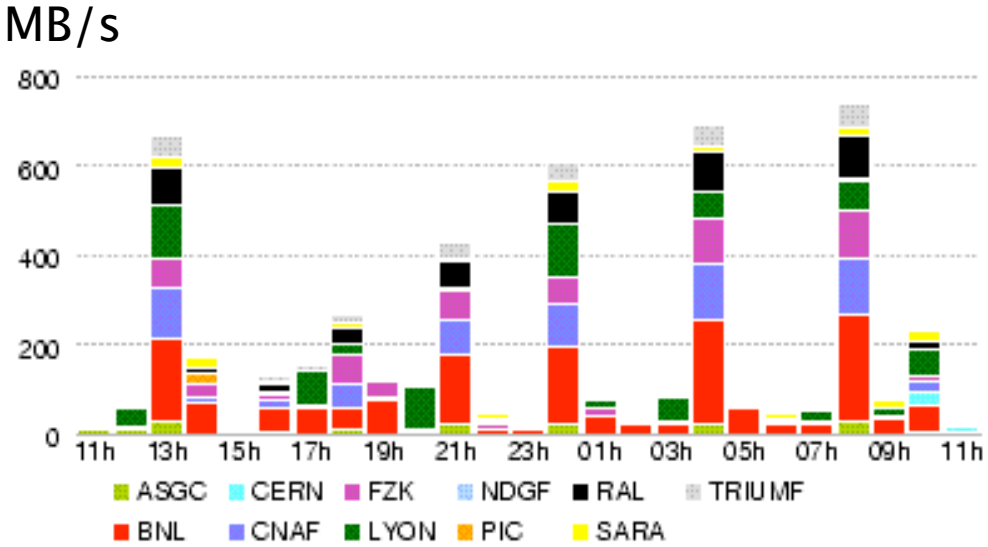
Problem at load generator on 27th

Power-cut on 30th

T1->T2 transfers

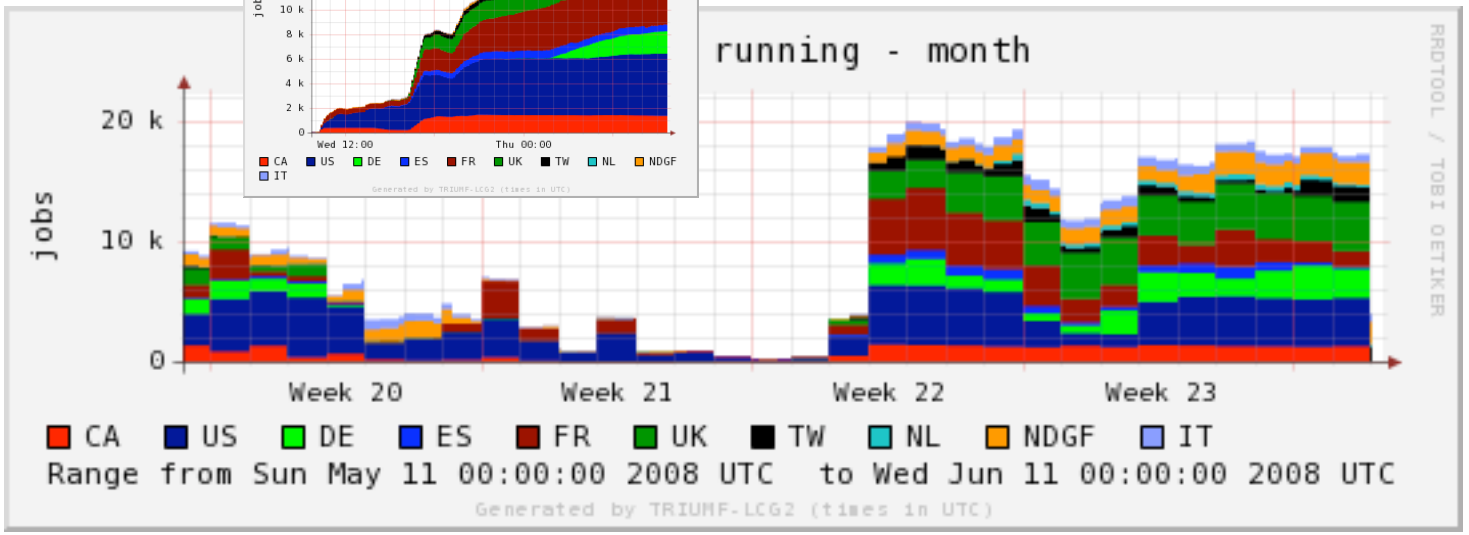
show a time structure

Datasets subscribed:
-upon completion at T1 -every 4 hours

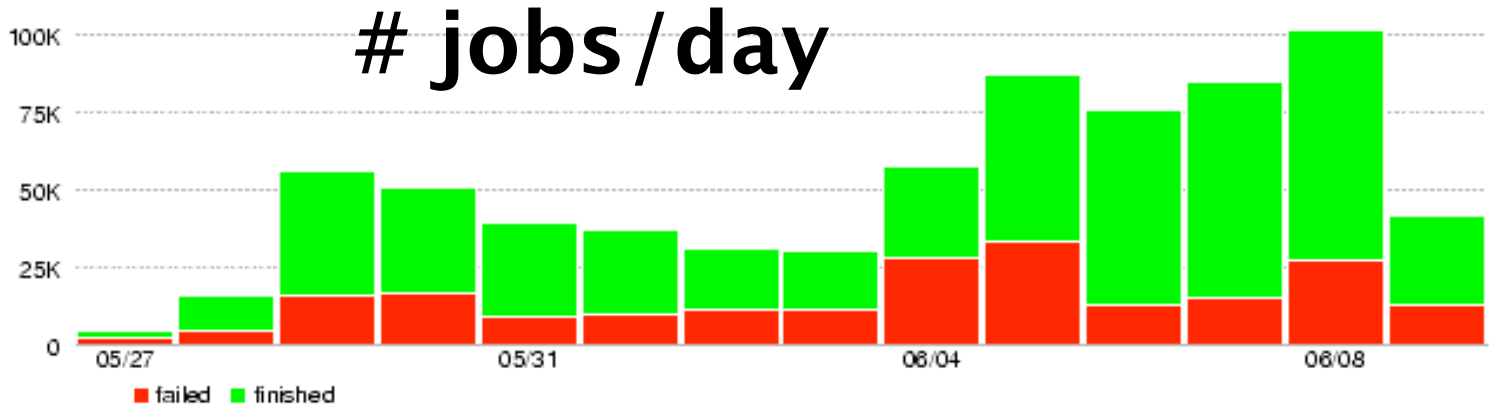


Week-4 and beyond: Production

running jobs

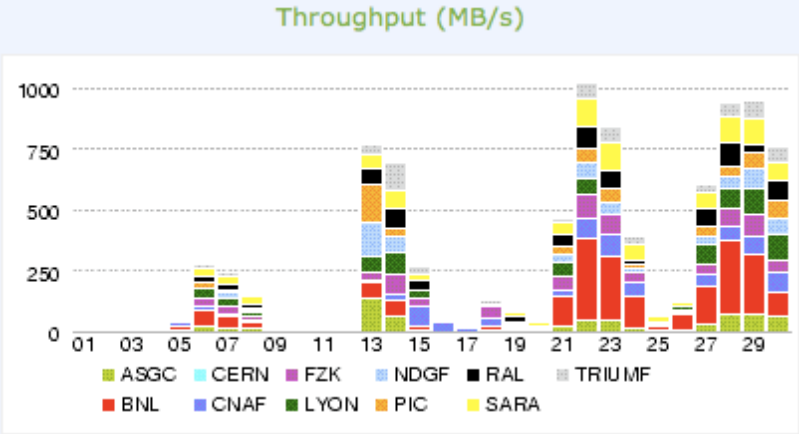


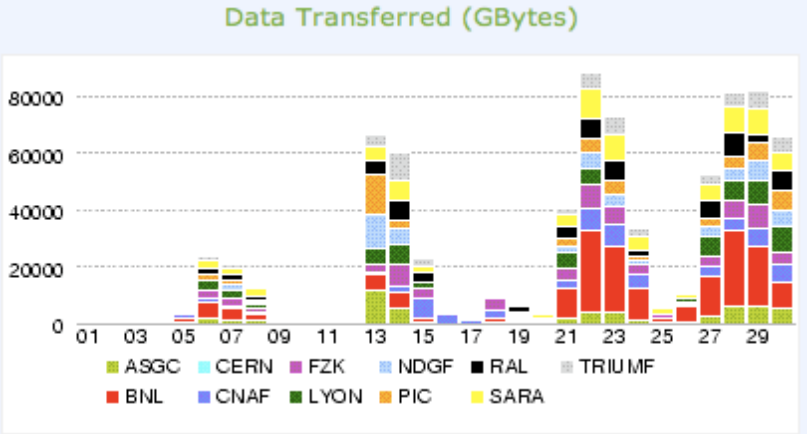
jobs/day



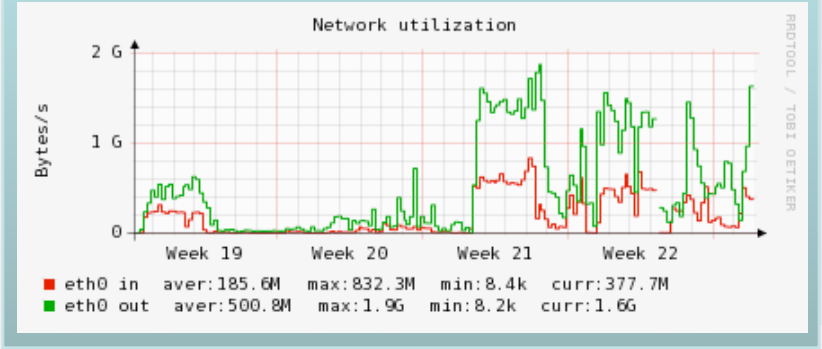
All month activity

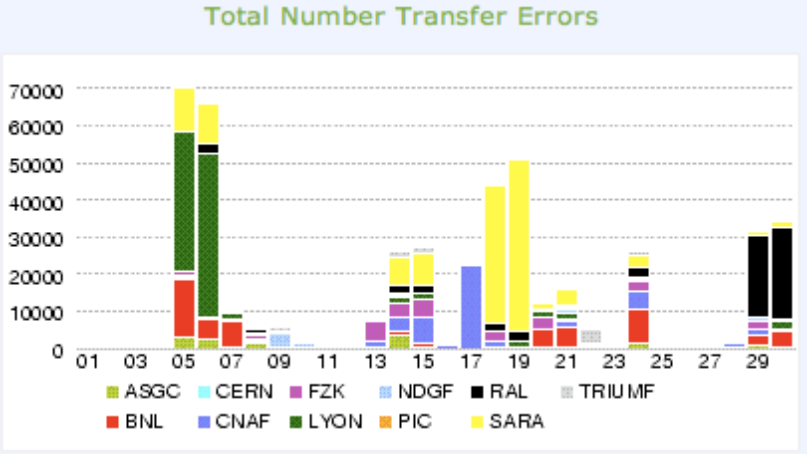
This includes both CCRC08 and detector commissioning



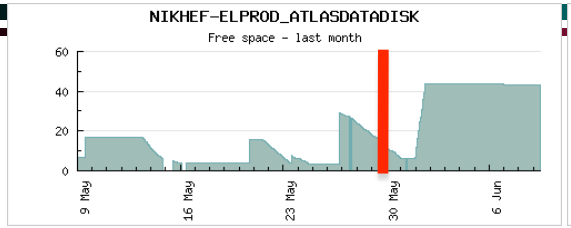


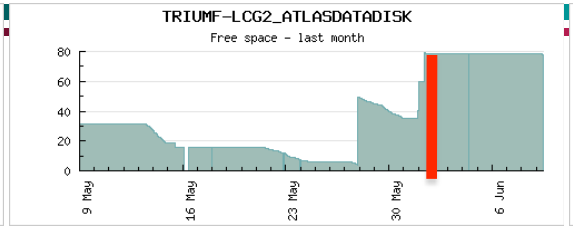
Heavy load activity on CASTOR@CERN

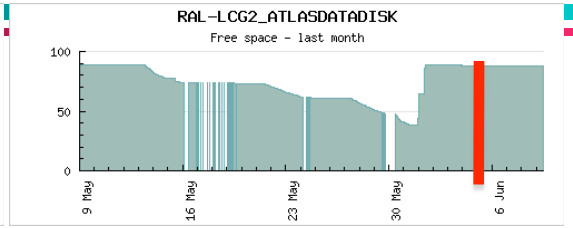


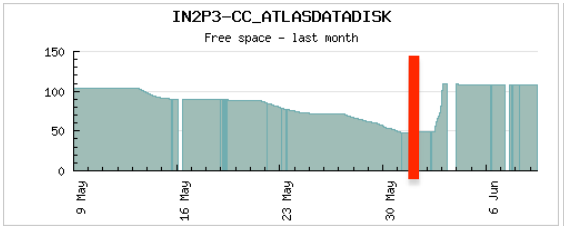


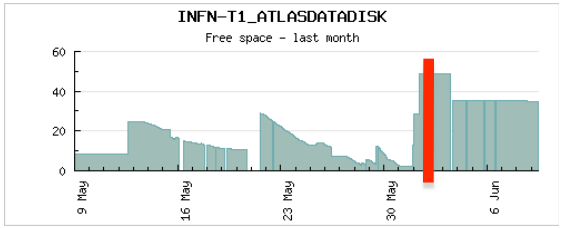
Disk Space (month)

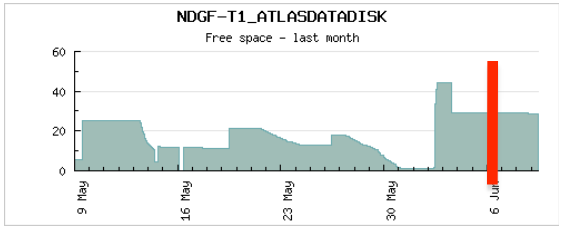


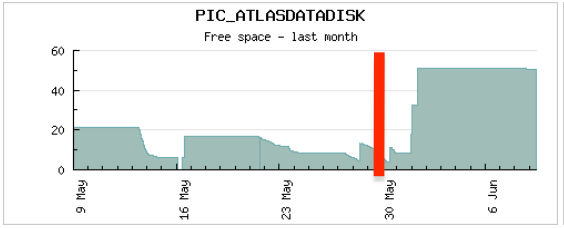


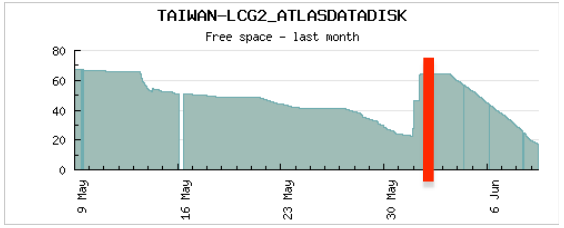


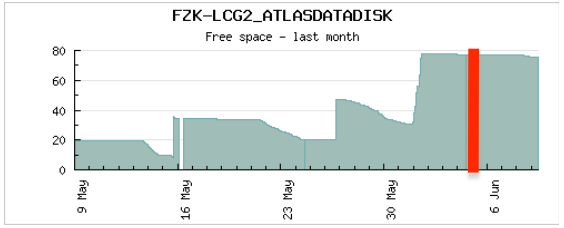


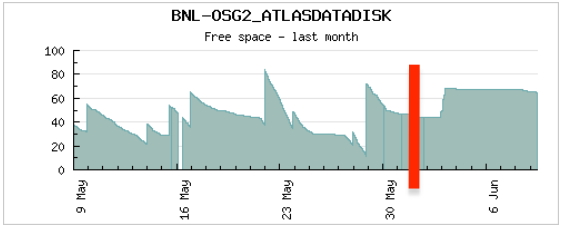












ATLAS "moved" **1.4PB** of data in May 2008

1PB deleted in EGEE+NDGF in << 1day

Possibly another **250TB deleted in OSG**

Deletion agent at work. Uses SRM+LFC bulk methods.

Deletion rate is more than good (but those were big files)

CCRC Conclusions

- The ATLAS data distribution scenario has been tested well beyond the use case for 2008 data taking
- The WLCG GRID infrastructure met the experiment's requirements for the CCRC'08 test cases
- Human attention is still needed
- Activity should not stop
  - ATLAS will now run continuous "heartbeat" transfer exercise to keep the system alive
  - At 10% bandwidth rate

ATLAS Readiness for Data Conclusions

- We've attempted to stress test all components of the software and computing system prior to exposing it to first physics data
  - A mix of component and large scale tests
  - Cosmics tests every ~month
    - Not discussed in this talk
  - Full Dress Rehearsal (FDR)
    - Large scale tests using simulated data samples using TDAQ and Tier-0 infrastructure
  - Combined Computing Readiness Challenge (CCRC)
    - Data Distribution rate tests run simultaneously with all LHC experiments
- We've put into place an incremental release building and patching infrastructure designed to allow fast run-around of validated patches in response to the unexpected
- Only time will tell whether we've anticipated everything that will go wrong!